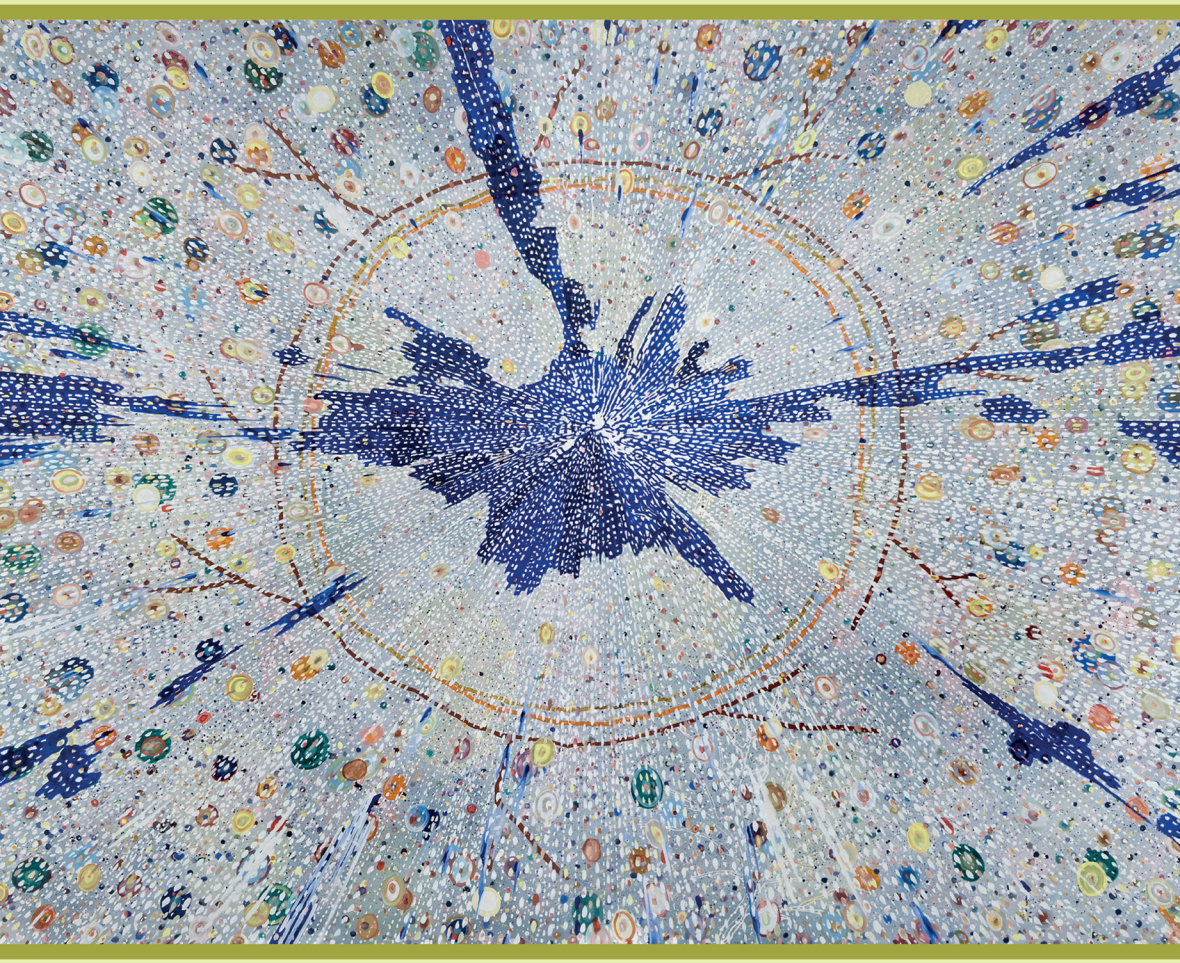
Data Politics

WORLDS, SUBJECTS, RIGHTS



Edited by
Didier Bigo, Engin Isin,
and Evelyn Ruppert

Routledge Studies in International Political Sociology

ROUTLEDGE

DATA POLITICS

Data has become a social and political issue because of its capacity to reconfigure relationships between states, subjects, and citizens. This book explores how data has acquired such an important capacity and examines how critical interventions in its uses in both theory and practice are possible.

Data and politics are now inseparable: data is not only shaping our social relations, preferences, and life chances but our very democracies. Expert international contributors consider political questions about data and the ways it provokes subjects to govern themselves by making rights claims. Concerned with the things (infrastructures of servers, devices, and cables) and language (code, programming, and algorithms) that make up cyberspace, this book demonstrates that without understanding these conditions of possibility it is impossible to intervene in or to shape data politics.

Aimed at academics and postgraduate students interested in political aspects of data, this volume will also be of interest to experts in the fields of internet studies, international studies, Big Data, digital social sciences, and humanities.

Didier Bigo is Professor of War Studies at King's College London and Research Professor at Sciences-Po, CERI Paris.

Engin Isin is Professor in International Politics at Queen Mary University of London, UK and University of London Institute in Paris (ULIP).

Evelyn Ruppert is Professor of Sociology at Goldsmiths, University of London.

Routledge Studies in International Political Sociology

Routledge Studies in International Political Sociology aims to provide a forum for outstanding empirical and theoretical research engaging with the interplays between the international, the political and the social. This timely book series draws upon significant theoretical and empirical challenges within the growing critical approach of international political sociology. It seeks to address, to encourage and to conceptualise the knowledge and understanding of transversal issues at stake when exploring the different components of the heterogeneous worlds hidden behind International Relations.

For more information about this series, please visit: https://www.routledge.com/Routledge-Studies-in-International-Political-Sociology/book-series/IPS

Perspectives From International Political Sociology
Transversal Lines in International Relations
Edited by Tugba Basaran, Didier Bigo, Emmanuel-Pierre Guittet & RBJ Walker

Data Politics
Worlds, Subjects, Rights
Edited by Didier Bigo, Engin Isin and Evelyn Ruppert

DATA POLITICS

Worlds, Subjects, Rights

Edited by Didier Bigo, Engin Isin, and Evelyn Ruppert

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2019
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
52 Vanderbilt Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Bigo, Didier, editor. | Isin, Engin F. (Engin Fahri), 1959- editor. | Ruppert, Evelyn Sharon, 1959- editor.
Title: Data politics : worlds, subjects, rights / edited by Didier Bigo, Engin Isin and Evelyn Ruppert.
Description: Abingdon, Oxon ; New York, NY : Routledge, 2019. | Series: Routledge studies in international political sociology | Includes bibliographical references and index.
Identifiers: LCCN 2018054225| ISBN 9781138053250 (hardback) | ISBN 9781138053267 (pbk.) | ISBN 9781315167305 (ebook)
Subjects: LCSH: Big data—Political aspects. | Big data—Social aspects.
Classification: LCC QA76.9.B45 D385 2019 | DDC 005.7—dc23
LC record available at https://lccn.loc.gov/2018054225

ISBN: 9781138053250 (hbk)
ISBN: 9781138053267 (pbk)
ISBN: 9781315167305 (ebk)

Typeset in Bembo
by Swales & Willis Ltd, Exeter, Devon, UK

CONTENTS

ILLUSTRATIONS

Figures

Table

Textbox

CONTRIBUTORS

The editors

Didier Bigo is Professor of War Studies at King's College London and Research Professor at Sciences-Po, CERI Paris. He is editor of the quarterly journal, *Cultures & Conflicts*, and was the founder and co-editor of the journal, *International Political Sociology*, published by the International Studies Association. His work concerns sociology of surveillance, policing, and borders. He co-edited *Transversal Lines* (with Tugba Basaran, Emmanuel-Pierre Guittet and R. B. J. Walker, 2016) as part of the *Routledge Studies in International Political Sociology*.

Engin Isin is Professor in International Politics at Queen Mary University of London and University of London Institute in Paris (ULIP). Isin's work concerns politics of the changing figure of the citizen as a political subject. He has authored *Cities Without Citizens* (1992), *Citizenship and Identity* (with Patricia Wood, 1999), *Being Political* (2002), *Citizens Without Frontiers* (2012), and *Being Digital Citizens* (with Evelyn Ruppert, 2015). He has edited *Acts of Citizenship* (2008) with Greg Nielsen, *Enacting European Citizenship* (2013) with Michael Saward and *Routledge Handbook of Global Citizenship Studies* (2014) with Peter Nyers. His latest book is *Citizenship after Orientalism: Transforming Political Theory* (2015).

Evelyn Ruppert is Professor of Sociology at Goldsmiths, University of London. She studies how digital technologies and the data they generate can powerfully shape and have consequences for how people are known and governed and how they understand themselves as political subjects, that is, citizens with rights to data. Evelyn is PI of an ERC funded project, Peopling Europe: How data make a people (ARITHMUS; 2014–19). She is Founding and Editor-in-Chief of the SAGE open access journal, *Big Data & Society*. Recent books are *Being Digital Citizens* (with Engin Isin, 2015) and *Modes of Knowing* (with John Law, 2016).

The contributors

David M. Berry, Professor of Digital Humanities, University of Sussex; Visiting Fellow, School of Advanced Studies, University of London; and Associate Member, Faculty of History, University of Oxford.

Laurent Bonelli, Associate Professor, Political Science, University of Paris-Nanterre.

Jonathan Cable, Lecturer, School of Media, University of Gloucestershire.

Ronald J. Deibert, Professor, Political Science and Director, Citizen Lab, Munk School of Global Affairs and Public Policy, University of Toronto.

Lina Dencik, Reader, School of Journalism, Media and Culture, Cardiff University.

Paul N. Edwards, William J. Perry Fellow in International Security, Center for International Security and Cooperation, Stanford University, and Professor of Information and History (Emeritus), University of Michigan.

Jennifer Gabrys, Professor, Department of Sociology, University of Cambridge.

Elspeth Guild, Jean Monnet Professor *ad personam* Queen Mary University of London and Radboud University Nijmegen.

Arne Hintz, Senior Lecturer, School of Journalism, Media and Culture, Cardiff University.

David Lyon, Director, Surveillance Studies Centre, Professor of Sociology and Professor of Law, Queen's University.

Brett Neilson, Professor, Institute for Culture and Society, Western Sydney University.

Louis W. Pauly, J. Stefan Dupré Distinguished Professor of Political Economy, Department of Political Science and Munk School of Global Affairs and Public Policy, University of Toronto.

Ned Rossiter, Professor of Communication, Institute for Culture and Society and the School of Humanities and Communication Arts, Western Sydney University.

Félix Tréguer, Postdoctoral Researcher, Centre National de la Recherche Scientifique (CNRS).

Tommaso Venturini, Advanced Research Fellow, French Institute for Research in Computer Science and Automation (INRIA).

Giovanni Ziccardi, Legal Informatics Chair and Director of the Information Society Law Centre (ISLC), Faculty of Law, University of Milan.

ACKNOWLEDGEMENTS

We would like to thank the editors of the series, *Routledge Studies in International Political Sociology* for their support and encouragement. Many thanks also go to Claire Maloney and Robert Sorsby at Routledge who helped guide this book through the production process. We would also like to thank Margaret Gillespie Cheesman, a PhD candidate at the Oxford Internet Institute, for the editorial assistance she provided. Her attentiveness and professionalism made invaluable contributions to the production of this collection. We are grateful for the contributions the authors made by responding to our initial framing and provocation and delivering inspiring chapters that enrich our understanding of data politics. We acknowledge that the research leading to this book was supported by funding from a European Research Council (ERC) Consolidator Grant 615588 (ARITHMUS; PI Evelyn Ruppert) and a French National Research Agency (ANR) Grant (2014-UTIC: PI Didier Bigo).

1

DATA POLITICS[1]

Didier Bigo, Engin Isin, and Evelyn Ruppert

Introduction

In 1983, Ian Hacking (2015) described the period between 1820 and 1840 as the "avalanche of printed numbers" in Europe and America. Hacking was reflecting on Michel Foucault's concept of biopolitics that targeted "population" with its own characteristics as an object of government in the nineteenth century. This invention was related to developments not least the birth of a science – statistics, Hacking's primary concern – but also associated sciences such as demography and probability, and data production practices such as the census and administrative registers. Hacking emphatically characterised that as the period when the "statistical study of populations comes to amass gigantic quantities of data" (2015, 280).

As Hacking was identifying "gigantic quantities of data" a new term was rapidly becoming popular in Euro-American languages: "personal computer". The invention of large-scale data processing machines following the Second World War was giving way to the miniaturisation of both processors and components of a computer – storage, graphics, controllers, and cooling. By the late 1980s a personal computer could already store and process all the "gigantic data" collected about populations between 1820 and 1840. This would have been truly wondrous to William Farr (1807–1883), a compiler of abstracts for the newly-founded (1836) Office of the Registrar-General of England and Wales who remained in office for 40 years (2015, 284). Hacking recounts that it was Farr who pirated a Swedish computing machine with more than 5,000 moving parts for use in the Office of the Registrar-General (Hacking 2015, 291). Two centuries later, a hand-held device could store and process such gigantic data.

Although the contemporary period has been described as the era of data revolution (Kitchin 2014, Mayer-Schönberger and Cukier 2013), we insist that it be

placed in a longer history. The personal computer of the 1980s morphed into a ubiquitous device of the twenty-first century, became connected with other devices through the Internet (a word born in 1980), converged storing and processing of data with sharing, led to the invention of protocols for collecting, representing, and sharing of data, and generated not only an Internet of people but also of things. Now, the amount of data generated and collected from these devices and the interests, authorities, and expertise required to render them useful make the data revolution of the 1820s appear rather miniscule but we need to understand the present as part of a broader historical transformation.

When Edward Snowden, a security operative working for the CIA, walked out of his office for the last time in 2013 (thereafter he became an exile), to reveal that national security organisations had been "harvesting" and "mining" gigantic masses of data generated by devices, he was carrying a small storage device capable of holding thousands times more data than was amassed between 1820 and 1840 (Bauman et al. 2014, Lyon 2014, Toxen 2014). His act revealed not only the truly enormous quantities of data that have been amassed from devices about those who use them and their interconnections and communications but also the varieties of analytical and algorithmic technologies invented to analyse and interpret them. The question now is how to place the 1980–2020 period within a broader historical transformation?

This book attempts to step back from these developments to position them within a broad historical-sociological perspective to articulate an international political sociology of data politics. We offer it not to express awe in contemporary technological developments but draw attention to social and political practices and arrangements that made them possible. Unlike many interpreters, Hacking understood Foucault's work as involving different histories of life, labour, and language and argued that Foucault provides both short and long histories of life (Hacking 2015, 279). He saw Foucault's distinction between body politics (discipline) and biopolitics (regulation) as different perspectives on the same series:

> There is a longer and a shorter story of biopolitics. The longer story gradually assumes a definite form in the mid-eighteenth century, and it continues today. Whereas Foucault's early books talked of sharp transformations, his research on sexuality directs itself not to mutation and revolution but to evolution in the longer term. There is no inconsistency in this: the world knows both revolution and evolution.
>
> *(Hacking 2015, 281)*

The sharp change that Hacking detects in 1820–1840, he argued, belongs to the longer story of biopolitics. Moreover, he also admitted that 1820 and 1840 are arbitrary dates and more precisely these should be 1839 and 1848. Why? The gigantic accumulation of numbers actually bracketed two European political revolutions:

> It represented an overt political response by the state. Find out more about your citizens, cried the conservative enthusiasts, and you will ameliorate their conditions, diminish their restlessness, and strengthen their character. Statistics, in that period, was called moral science: its aim was information about and control of the moral tenor of the population.
>
> *(2015, 281)*

Hacking then goes on to illustrate how calculating machines originated from the need to collect, store, and analyse these numbers and how the longer history of biopolitics made the conditions of possibility of the invention of statistics as a moral science of the state and how this science has driven calculating machine technologies in the late-nineteenth and early-twentieth centuries. Although there have been various studies since Hacking's article that explored the rise of census, survey, and statistical technologies as developments of biopolitics (Desrosières 1998, Porter 1986), we want to see recent developments within a similar series.

The purpose of this book is to think about recent transformations in data politics. In our introduction we position these in historical-sociological terms, especially of the kind that Foucault and Bourdieu initiated and Hacking and others expanded and modified. For there are fundamental differences between empire-states amassing gigantic amounts of data for governing metropole and colonial populations in the nineteenth century and the complex assemblage of public and private authorities and interests invested in the production of data in the twenty-first century. This book is certainly about these differences. But it is also about situating these differences in relation to social, economic, and political conditions when such a modern regime of government emerged and of which we are still subjects. As a contribution to international political sociology we want to consider the conditions of possibility of data politics as a field of power and knowledge (Bigo 2011, Bigo and Walker 2007, Bonditti, Bigo, and Gros 2017).

What is data politics?

If not for the rapid development of the Internet and its connected devices "data" would have probably remained a relatively obscure concept or term confined to these sciences. Yet, data has become a social and political issue not only because it concerns anyone who is connected to the Internet but also because it reconfigures relationships between states, subjects, and citizens. Just about every device is now connected to the Internet and generating vast quantities of digital traces about interactions, transactions, and movements whether users are aware or not. What started as an ostensibly liberated space rapidly became the space over and through which governments and corporations began collecting, storing, retrieving, analysing, and presenting data that records what people do and say on the Internet. This ranges from who communicates with whom, who goes where, and who says what – and much more besides. This is now being augmented with data that people collect

about themselves, especially their relations, body movements, and measurements; the amount and range of data that has become available is, as everyone now knows, staggering. There has never been a state, monarchy, kingdom, empire, government, or corporation in history that has had command over such granular, immediate, varied, and detailed data about subjects and objects that concern them. What exactly governments, corporations, and a whole series of agencies and authorities collect, analyse, and deploy is complex but it is now generally understood that data has become a major object of economic, political, and social investment for governing subjects. This development has been captured by the term "big data" to mark a departure from conventional forms of data and statistical knowledge. While first coined by industry, big data has come to have different meanings and uses but significantly, and along with the increasing ubiquity of data in everyday life, the term has become less prominent. Notably, attention has started shift to a focus on computation and analytics such as algorithms, machine learning, artificial intelligence, and the Internet of things. Yet, data remains a key matter of concern as both the product and condition of computation and analytics.

Scholarship on these developments has understandably focused on issues concerning surveillance, privacy, anonymity, and types of conduct that the Internet cultivates about always-connected, always-measured selves. Perhaps equal to the measure of the influence of the Internet there has been scholarship on data ranging from warnings about its consequences (surveillance, privacy, isolation) to types of conduct (racism, misogyny, bullyism). Along with this, numerous studies, reports, guidelines, regulations, and legislation concerning data protection and the rights of data subjects have proliferated.

Data Politics builds on this scholarship but it aims to make three distinct yet interrelated contributions to an international political sociology of data politics.

The first concerns a shift in focus from the politics of or in data to data as a force that is generative of politics. In this view, rather than settled in databases or archives, data is a force realised through its production, uptake and deployments. We want to draw the implications of thinking about data not as an inert representation but a language with performative force as Bourdieu (1993, 1973) and Butler (1997) have shown. That is, data politics is concerned with not only political struggles over data production and its deployments, but how data is generative of new forms of power relations and politics at different and interconnected scales. If indeed data enacts that which it represents, this signifies two things. To collect, store, retrieve, analyse, and present data through various methods means to bring those objects and subjects that data speaks of into being. Data sciences such as statistics, probability, and analytics have emerged not because they have merely quenched our curiosities but because these sciences have been useful for the objects and subjects they have brought into being for the purposes of governing and/or profit. And, to speak constantly about data as though it either represents or records subjects and objects and their movements, independent from the social and political struggles that govern them, is to mask such struggles.

That data is generative of new power relations and politics is evident in the recent struggles over how big data was allegedly used in the US election and UK

referendum to create personalised political advertising to influence how people voted. Referring to these electoral uses, George Monbiot writing in *The Guardian* noted that we must act now to own these new political technologies before they own us. He was of course referring to the work of a company called Cambridge Analytica, which was partly owned by US billionaire Robert Mercer, who also happens to be a friend of former UKIP leader Nigel Farage. It was widely reported that the company allegedly influenced both the US election and the UK referendum by mining data from Facebook and using it to create profiles predicting people's personalities and then tailoring advertising to their psychological profiles. While some of the claims that this happened were brought into question, including denials from Cambridge Analytica, the UK's privacy watchdog – the Information Commissioner's Office – deemed there was sufficient cause to launch an inquiry. These claims and denials were soon followed by the disclosure that the personal information of up to 87 million users was harvested without their permission by an app designed by a Cambridge academic. The seriousness of this breach intensified when Cambridge Analytica claimed that hundreds of companies harvest such data and that it is legal to do so. Or when the Cambridge academic at the centre of the controversy claimed that it was both legal and ethically acceptable to sell data to a third party. Or when CEO Mark Zuckerberg admitted that Facebook took no action to ensure that the tens of thousands of apps it approved adhered to their terms of service.

So, in the wake of already uneven power and influence over electoral processes – such as campaign financing and media alliances – we now have misinformation, disinformation, and techniques such as bot-swarming whereby fake online accounts are created to give the impression that large numbers of people support a political position. For these reasons, Oscar Gandy recently argued that this calls for a shift of attention away from a focus on privacy or surveillance and the collection and processing of information to how information is being used and misused (Gandy and Tsui 2018).

What these examples illustrate is that data and politics are inseparable. Data is not only shaping our social relations, preferences, and life chances but our very democracies. And that is how we want to speak of data politics. However, a problem with these views on data politics is that the subjects who are constituted as the addressee are presumably the affected Internet subjects. This is the second intervention that has led us to articulate what we call data politics. It concerns atomism: often such pronouncements address atomised individuals who need to protect themselves from the dangers of the Internet and its manipulations. It is based on the ontological premise of "hyper-individualism" whereby persons, events and phenomena are treated as independent and "atomistic" entities (Lake 2017). Data politics that emerges from this reaction is one of urging people to protect themselves as individuals. It is almost as if the narrative says "yes, there is collective work that needs to be done but ultimately it is up to you to change your behaviour to protect yourself from the dark forces of the Internet". The addressee in other words is the atomised subject whose data is individualised rather than

understood as a product of collective relations with other subjects and technologies (Socialising Big Data Project 2015).

A third intervention concerns the immediacy that pervades these reactions or responses. They are predominantly exercised by the immediacy of a threat, danger, menace, risk, or peril or insecurity or unease that the Internet ostensibly engenders. Even those who have fought battles with governments and corporations to expose their data practices fall prey to a Messianic creep in articulating political problems by decrying their immediacy.

The obverse response to these reactions has been to extol the virtues of the Internet and illustrate that if it is not liberating it is at least making our lives better organised, measured, improved, whatever. Yes, there may be dangers and insecurities but this is a small price to pay for the benefits it brings. This response is still riddled with immediacy and atomism. Its calculative logic is from the point of view of the atomised subject weighing the pros and cons of the Internet against the threats of immediacy.

All this has led us to the conclusion that data politics is yet to find its subjects. This book attempts to step back from the inertness, atomism, and immediacy of the dominant points of view of the Internet and the data it generates and ask questions about data politics and position these within a broad historical-sociological perspective. What do we then mean by an international political sociology of data politics?

We start with the assumption that the will to knowledge and the will to power are two aspects of how we conduct ourselves and the conduct of others, and thus we approach data not as a representation (i.e., information collected, stored, and presented without interest) but as an object whose production interests those who exercise power. This was at least one of the lessons we have learned from Michel Foucault's studies of the ways in which modern societies come to depend on governing subjects with data collected over not only their physical and social attributes (life, language, labour) but also about the conduct of their behaviour (Foucault 2007). Our second assumption is that the production of data is a social and often political practice that mobilises agents who are not only objects of data (about whom data is produced) but that they are also subjects of data (those whose engagement drives how data is produced). Our question thus shifts to social practices and agents. Just as the avalanche of numbers was an aspect of the birth of a modern regime of government, in our age data does not happen through unstructured social practices but through structured and structuring fields in and through which various agents and their interests generate forms of expertise, interpretation, concepts, and methods that collectively function as fields of power and knowledge. This was at least one of the lessons we learned from Pierre Bourdieu's studies on the ways in which fields of knowledge constitute fields of power (Bourdieu 1988) that involve struggle and change, fragile moments, and the emergence of new kinds of practices (Bigo 2011).

Foucault and Bourdieu influenced a generation of scholars who have taken up the relations between power, knowledge, and fields and investigated the ways in which states, agencies, organisations, corporations, and institutions – often assembled

in different combinations as governments – constituted their authority, legitimacy, and legality by producing knowledge about objects and subjects through establishing method and data regimes such as censuses, indexes, indicators, registers, rolls, catalogues, logs, and archives. We now understand much better the relationships between state formation and statistics, probability, and data regimes (Desrosières 1998, Hacking 1990, Porter 1986). Statistics, from their very beginning, combined "the norms of the scientific world with those of the modern, rational state" (Desrosières 1998). These data regimes have now been extensively studied as historical developments. The birth of objects of knowledge such as the economy, population, society and their sciences – originally called political arithmetic and now statistics — have also been studied extensively. Although it would be impossible to summarise what we now know about these data regimes and the state, the overall insight we have gained can be stated as follows. While Max Weber's argument that the sovereignty of the state consists in its monopoly of the means of violence is often cited, following the studies of Foucault and Bourdieu and the literature inspired by them, we have come to recognise that this sovereignty depends on numerous practices beyond the organisation of violence. Historically, the state performs sovereignty with control over and dependence on especially education, fiscal, and cultural data regimes. This does not mean that citizens in each state did not influence, interfere, or intervene in the ways that data regimes constituted them as data subjects. On the contrary, scholars have also investigated and documented how citizens have developed democratic practices to challenge social categories of data regimes and their effects (Anderson and Fienberg 2000, Kertzer and Arel 2002, Nobles 2000). There are many cases that illustrate how, for example, census categories such as race, ethnicity, gender, and other indexes have been called into question, subverted, and transformed.

Nonetheless, the state, or rather organisations, institutions, agencies, agents, and authorities that make up the complex field of government, maintained an effective monopoly on data regimes concerning whole populations. This is not to say that corporations did not also generate data about their customers especially over the last century or so but this was largely limited to specific population groups and in relation to narrow concerns. Beginning in the early-twentieth century, opinion polling and marketing research were considerable developments in corporate forms of population data generation (Osborne and Rose 1999). And although there have been various international organisations that have entered into fields of data generation and accumulation such as the United Nations, the European Union, Organisation for Economic Co-operation Development, and the International Labour Organisation, the primary site and scene of the collection of population data and its various regimes have remained the monopoly of the state for nearly four centuries.

This monopoly of the state over data production, collection, and even interception is increasingly challenged. Or, at least, state sovereignty over data regimes is now shared by the birth of entirely new assemblages of the production of data (Kitchin 2014). Not least has been the increasing accumulation and mobilisation of data by corporations (Thrift 2005). It is tempting to immediately single out the Internet and its connected devices as the source of this challenge. But it is

much more complicated than that as our argument above anticipates. It would be folly to assume that Internet technologies develop independently from the interests that constitute the fields through which various data regimes have been invented. However, beyond technological developments, the sovereignty of the state in accumulating and producing data about its population, territory, health, wealth, and security is being challenged by corporations, agencies, authorities, and organisations that are producing myriad data about subjects whose interactions, transactions, and movements traverse borders of states in new and complicated patterns. Not least, these traversals challenge the methodological nationalism that has dominated statistical thought and practice and their corresponding boundaries of population data, knowledge and power for centuries (Scheel et al. 2016). While Bourdieu's studies focused on the nation and in particular France, others have taken up his conception to understand fields as international and transnational (Dezalay and Garth 1996, Madsen 2011, 2014). For Bigo, the transnational exists in the form of transnational networks and practices of professionals who "play simultaneously in domestic and transnational fields" (Bigo 2011). In this view, a transnational field is constituted by networks and practices between and amongst professionals who act at various non-hierarchically ordered scales of the transnational, national, and local (Scheel et al. 2016).

We have divided the book into what we consider as three domains of data politics: worlds, subjects, rights. In the first part, we discuss some key conditions of possibility of these domains of data politics and then in the next three parts the importance of each domain. We pose key questions that are not exhaustive of possible inquiries and then summarise the contributions of each chapter. Taken together, the chapters of this book set out political questions about the ways in which data has been constituted as an object vested with certain powers, influence, and rationalities.

Part I: conditions of possibility of data politics

Part I addresses some of the conditions of possibility of data politics and through which new worlds are produced, new subjects come into being, and new rights emerge from struggles over the ownership, collection, analysis, and storage of data. The chapters in this part reveal some of the complexities of these conditions. In Chapter 2 Paul N. Edwards examines the role of infrastructures as one condition of possibility and specifically those of environmental data systems that have been built over a long period of time and are now being undermined by the Trump administration's attack on climate science in the USA. He demonstrates how data analysis models (or algorithms) that mine, collate, organise, and present data and their interoperability and compatibility have become infrastructures of knowledge about the earth's climate. These data models he argues have now become primary worlds of struggle over knowledge about climate change. He deftly illustrates the tension between critiques of algorithms that critical data scientists advance and the consequences of eliminating such data models as infrastructures of knowledge.

In Chapter 3 David M. Berry thinks "beyond data" to critically consider their algorithmic underpinnings and connections to a wider political economy and across multiple levels of computational systems. He examines the complexity of understanding the code that underlies data models or algorithms. Berry points out a paradox of the Internet where billions of people communicate on the basis of a language that is hardly visible or comprehensible to them: the code. So, while the Internet may depend on a massive infrastructure of servers, devices, and cables what brings them together or more precisely what holds them together and enables them to communicate with each other is this special kind of language. But to understand code is anything but straightforward because code itself embodies various programming and communication languages such as binary machine code to algorithms (Galloway 2006). The Internet has a language but it is hardly visible or even comprehensible to those who do not write such code. How does the language of the Internet traverse both actual and virtual worlds of data? Berry argues that the struggles over the language of the Internet and its code takes place simultaneously with the struggles over "natural" languages and their use and abuse. The question then becomes to what extent those who write code enable and shape the former. In regards to this question he argues for a critical theory of algorithms (CTA) to examine "the particular historical conditions that give the present its shape in relation to the specific material and ideological formations that algorithms introduce into the social and economic conditions of society."

In Chapter 4 David Lyon focuses on how everyday life in the twenty-first century is unavoidably surveillant, especially in the increasingly data-dependent Global North and South. This condition is led by giant Internet corporations such as Google who promote data capture and analysis as the new fuel for prosperity and progress, which raises profound questions for the politics of data and everyday life. Lyon frames his discussion of this condition in terms of two wide-ranging concepts, surveillance capitalism and surveillance culture, which both depend on data but often in different ways and with different consequences. He argues that surveillance capitalism is the source of systems that enable many aspects of surveillance culture, and that at present much that counts as surveillance culture is supportive of surveillance capitalism. But, he contends, this is not inevitable, as evident in the case of the Facebook scandal of 2018. The conditions of possibility – surveillance data in this case – do not produce predetermined outcomes. Instead, Lyon argues that a meaningful data politics can emerge through the reassertion of human dignity and especially agency in responses to surveillance capitalism.

Each of the following parts of the book provide more detailed investigations of the worlds, subjects, and rights that emerge under these conditions of possibility of data politics.

Part II: worlds

The Internet is an elaborate infrastructure composed of objects, equipment, cables, routers, servers, switches and devices that constitute a unique technological materiality. Unlike other massive material transformations of industrial and post-industrial

cities and their transportation and communication infrastructures, the materiality of the Internet is mostly out of sight and located elsewhere. The data servers and data farms are often in faraway and remote locations or nestled within cities that are inaccessible and unknown to most people. Its connectors are often buried under the earth or sea. Its wireless communications are invisible but routers, switches, and masts create strange yet recognisable objects within and outside cities. Without this massive infrastructure and its maintenance and production the Internet of things, communications, and exchanges would be impossible. The material infrastructure of the Internet not only generates new logics of borders and capacities of control that remain often invisible but also protocols and platforms that make people think the Internet is made up of a seamless and invisible flow of information. How are these worlds created and governed? What are the material conditions of possibility, configurations, and stratifications of these worlds? How do these worlds straddle or cross between offline and online worlds? To think of worlds is to trace how material conditions of the Internet are critical infrastructures that are generative of politics and struggles.

Through the Internet a new space is being made – a cyberspace perhaps – but understanding this space is fraught with difficulties. The Internet has not only blurred the boundaries between online and offline worlds but it has also rendered the distinction between the two spurious and perhaps untenable. With always-connected devices it is impossible to say when people or things are offline or online or indeed to separate embodied subjects from their operation. What kind of space does the Internet generate? What is the role of data in such a space and how does data make it possible? In turn, how does the Internet and the space it generates make data politics possible and with what effects? In Chapter 5 Ronald J. Deibert and Louis W. Pauly take up some of these questions by illustrating how states have been attempting to impose their borders on cyberspace. The expansion and intensification of controls over cyberspace by states within conventionally conceived territorial boundaries are well known. But they argue that states simultaneously project power in and through global cyberspace outside of their territorial jurisdictions. They remind us that struggles over cyberspace do not stop at borders and that extraterritorial projections of state power through cyberspace are expanding, deepening, and becoming more elaborate. They create a sophisticated image of cyberspace as a site of international politics and struggles between various national and international authorities.

The emergence of big data with its focus on production, accumulation, mining, circulation, aggregation, analysis, and interpretation has also engendered the formation of various professions from data scientists to data journalists. Each of these professions is engaged in competitive struggles between each other and with other professions and yet at the same time also reinforce the broader practice of investing data with powers. These emergent professions and their practices have not only begun reorganising existing fields of data production such as the official statistics of states (state and statistics share common etymologies) but also have given birth to new forms of data accumulation and valuation whose source of authority and legitimacy traverse the boundaries of state sovereignties and produce international effects.

In this light, data is not an already given artefact that exists (which then needs to be mined, analysed, brokered) but an object of investment (in the broadest sense) that is produced by the competitive struggles of professionals who claim stakes in its meaning and functioning. They engage in struggles over the valuation of different forms of capital conceived by Bourdieu including cultural, economic, social, and symbolic capital (Bigo 2013). It is through the accumulation of these various forms of capital that their relative positions are established within the field (Bourdieu and Wacquant 1992). The emergence of data as a field and data professionals as its custodians and gatekeepers shapes competitive struggles not only in defining an object but also the principles of how to understand and intervene in data politics. At the same time, algorithms increasingly call into question the very expertise that data accumulation has spawned through the automating practices of judgement. Who decides whether to invest, what to listen to, where to eat, where to stay, and where to go? How do algorithms embed expert judgements and normative assumptions without appearing to do so? In Chapter 6 Didier Bigo and Laurent Bonelli examine these issues through their analysis of the emergence of data production as a field and intelligence professionals as its producers. They argue that competitive struggles not only shape the defining of data as an object but also the principles of how to understand and intervene in what we call data politics. Through a Bourdieusian, international political sociology-inspired analysis, they illustrate the emergence of a transnational space where the production of security data occurs to argue against the illusionary idea of the intelligence community as a single world united by common surveillance techniques which are changing the understanding of security. Rather, they highlight how logics of action cut across and transgress distinctions between the internal and the external, the national and the foreigner.

The accumulation of data procures not only cultural capital but also economic capital. An economy of data is founded on the "voluntary" input of personal data in exchange for Internet services. This creates the conditions for the making of a stock market of data involving data brokers and profit shares generated by deep data mining and data discoveries. How do individuals contribute to this production and what is the political economy of desire that generates a material economy of services? What are the consequences of subjects giving up data in return for so-called free services? What are the legal conditions that enable and disable the circulation of data within and across states? From questions of data commons to data ownership, how are legal regimes being challenged and remade by struggles over data as property?

In Chapter 7 Tommaso Venturini takes up some of these questions through a focus on what is at stake with "fake news" as a key object of data politics. He illustrates that this misleading term conceals that the production of news and production of truth in general always involve interpretive struggles and a competition between interests to establish authority and expertise. Rather than considering it an object of algorithmic intelligence, computational analytics or political intentions, he proposes an understanding of fake news based on its circulation rather than its contents. He proposes that it is more appropriate to consider circulating stories as "junk news" and describes its economic, communicational, technological, cultural,

and political dimensions. In this way, Venturini shifts attention to our ability to discern between news and junk as an important object of debate and discussion and form of data politics.

Félix Tréguer in Chapter 8 considers how data politics is embodied in security assemblages – combinations of technology companies and security professionals – and how their practices are increasingly shaping how the state governs its citizens. He illustrates how these assemblages are leading to a new technological bureaucratisation of the state that transforms citizens' understanding of themselves as subjects of government. His chapter identifies the need to resist the technological bureaucratisation of the state as a significant element of data politics today – a theme that is picked up by chapters in Parts III and IV.

Part III: subjects

The emergence of data as an object of government engenders the emergence of subjects who take positions in and through the various resignifications and challenges that it spawns. Rather than occupying already existing positions, subjects are produced through various digital interactions and at the same time their digital traces shape and organise their subjectivities and how they are known and governed. How are subjects part of the work and making of data through which they then come to be known? Through procedures of channelling, filtering and sorting data, various devices and platforms configure not only transactions and interactions but the data they generate recursively shapes and forms subjects in never fixed but modulating ways. With the increasing circulation, mining and combining of data how are subjects and their affiliations, connections and relations multiplied and governed via ever more dispersed micro data politics?

People govern their health by making themselves data subjects of health. Measuring their own performances with Internet-enabled devices and benchmarking their performance against other performers, data subjects of health increasingly calibrate a model body not through images circulated by the advertising industry but by literally working themselves out through their data performances and for others. How is data part of the making and shaping of bodies and the body a site of data politics? Being a data subject entails the radically shifting meaning of being a consumer from a subject making choices to a choice-making and sorted subject. Being constantly a reviewer, modern consuming data subjects are caught in a spiral of evaluations: they are evaluated and evaluator all at once and all the time. Recommender platforms and evaluation data generated by transacting ever more sort subjects into categories of cultural preferences that narrow and channel choices. How is consuming through the Internet generative of data politics?

In Chapter 9 Lina Dencik, Arne Hintz and Jonathan Cable consider some of these questions of the data subject in relation to the uneven effects of data-driven surveillance practices which simultaneously advance particular social, economic and political agendas that enfranchise some whilst disenfranchising others, and prioritise certain ways of organising society at the expense of others. It is in relation to such

concerns that they consider the possibility of data justice. They note that much resistance to surveillance has predominantly centred on techno-legal responses relating to the development and use of encryption and policy advocacy around privacy and data protection. They argue that data surveillance should be considered in relation to broader social justice politics. If there is an emergent surveillance capitalism in which the collection, use and analysis of our data increasingly comes to shape the opportunities and possibilities available to us then we must ask broader questions of data justice.

The practices that produce data subjects also involve changing relations of production in the generation of data including the production of its labourers. Are we moving from the logic of having a job to a logic of contributing something to the fulfilment of a task? The data-generated market of global tasks has now created a vast meeting place for those who need and will pay for accomplishing specific and often micro tasks and those who can and need to fulfil these tasks to make a living. To consider the data subject also calls upon consideration of the uncanny convergence between robots and humans not in the way in which the cyborg manifesto (Haraway 1991) envisaged it but perhaps more in the manner in which Star Trek anticipated. How does the automated generation and analysis of data based on artificial intelligence and machine learning appear autonomous and yet inseparable from struggles and relations between programmers, subjects and technologies? In Chapter 10 Brett Neilson and Ned Rossiter approach these questions through the examination of data centres as sites of data politics. They show how data centres are increasingly moving toward automated economies with the integration of artificial intelligence, machine learning, and robotics into processes of capital accumulation. These data infrastructures should be considered sites of struggle not only because of where they are located but also how they have become hubs of command and control over production, consumption, and exchange circuits. Understanding how these centres regulate logistics by which various forms of capital is accumulated and how labour transitions to a society of automation for them is a key political question and field of struggle. For them, "data politics are not exclusive to the claiming of rights so much as the production of subjectivity within environments whose data architectures register conflicts between the politics of decentralisation-centralisation and the impossibility of pure distribution".

Data not only captures but also colonises minds, souls, bodies, and spaces. It subjectifies through practices of production, accumulation, aggregation, circulation, valuation, and interpretation. These practices call upon subjects who are not separate from but submit to and are active in the various ways that data is made and colonises lifeworlds to constitute "data's empire". In Chapter 11, Engin Isin and Evelyn Ruppert examine the various ways that data captures and colonises minds, souls, bodies and spaces and makes data subjects through practices of production, accumulation, aggregation, circulation, valuation, and interpretation. They draw our attention to the fact that these practices operate together yet differently in the metropole and postcolony and produce different data subjects. They remind us how European empires in the nineteenth century invented various data collection

and analysis methods for producing colonial populations and how contemporary practices build on these imperial infrastructures and logics. They invite readers to understand developments such as UN Global Pulse as instances of postcolonial data politics, which call for decolonising data politics.

Part IV: rights

If the accumulation of data traverses subjects it also constitutes them with claims to certain rights that concern its accumulation: who owns, distributes, sells, accesses, uses, appropriates, modifies, and signifies data become objects of struggles for claiming rights to such modalities. The rights claiming subject is the figure of the citizen that we have inherited as a political subject who is now making rights claims about being a subject of data. How do subjects exercise and claim such rights through what they say and do through the Internet? How do they perform rights and claims about being subjects of data through how they communicate, share, express, and engage with digital devices and platforms? How do they invent data practices that challenge and subvert state and corporate forms of data and struggle for rights through legal and regulatory mechanisms?

This third condition of data politics considers rights claiming subjects such as citizen data scientists as part of material-political arrangements and struggles over who generates, legitimises and has authority over data and how data is mobilised to make claims for environmental and other rights. It concerns how citizens make data an object of transnational politics and engage in struggles around free expression, privacy and ethics and the forums, practices, and networks through which these struggles are being fought. In Chapter 12 Giovanni Ziccardi shifts our attention from the collection and collation of data to consider rights over its life and death. He discusses the whole "life cycle" of data, especially from a legal-informatics point of view and with particular attention to the right to oblivion after the death of a person and how this constitutes a different kind of right. He discusses the complexities of the European Union General Data Protection Regulation (GDPR) and the impossibility of data oblivion. Rather, he argues that the right to data oblivion requires simultaneously addressing three forms of oblivion that make it up: social, which concerns the persistence and circulation of personal data; technical, which relates to the resistance of technology to the removal of data; and legal, which refers to forgetting, deleting, and de-indexing elaborated by legal means through case law or norms.

How are rights not only claimed through regulations, laws, and protocols but by citizens who make claims and in turn perform what is data politics through their everyday digital acts? In Chapter 13 Jennifer Gabrys takes up this question by shifting our attention from data as something collected about citizens to many instances where citizens generate their own data. Whether to document lived experiences through social media platforms, sensing air pollution to challenge governmental measurements, or documenting conflict in overlooked zones, citizens are collecting, analysing and communicating data to articulate alternative narratives. These practices of data citizens not only challenge official practices for making evidence, they also potentially reinvent

how rights are formed, expressed, and transformed through ongoing data practices. Gabrys show how citizen practices of using low-cost and digital sensor technologies to monitor air quality and changing urban environments generate distinct forms of data politics through the operationalisation of new data and data relations.

The relationship between the right to privacy and that of data protection is illustrative of the transversal relations and legal and political tensions that make up data politics. On the one hand, international human rights laws and obligations seek to secure and universalise the former and various national regimes have emerged to address the latter. However, transversal relations call for a figure of a citizen that is different from the subject we have inherited and instead one who can make rights claims that traverse national borders (Isin and Ruppert 2015). In Chapter 14 Elspeth Guild illustrates an emerging field of international law where data citizens are able to command and have control over their privacy. Guild notes that citizens have discovered to their shock how little control their own state authorities have over the protection of their privacy and shows that the global movement of communications, Internet, and social media platforms makes a citizen's right to privacy impossible to regulate and protect at the national level. Guild documents how since 2013 a number of authorities, interests, and forces have come together to create an international framework for privacy in a digital age. It is a framework that is emerging as a consequence of data citizens contesting and seeking to establish their rights to privacy by using the intersection of international and national law as a nexus through which to achieve their claims.

Conclusion

This book invites readers to regard contemporary transformations as a field of power and knowledge and an emerging regime of government that is comparable yet irreducible to the modern regime of government that emerged in the nineteenth century and of which we are still subject. It provides an analytical framing with a focus on worlds, subjects, and rights as conditions of possibility of such a field. Our hope is that the book contributes to our understanding of this field and the possibilities of data subjects becoming data citizens.

Note

1 An earlier version of this chapter appeared as a commentary and invitation to contributors to this book: (Ruppert, Isin, and Bigo 2017).

References

Anderson, Margo, and Stephen E. Fienberg. 2000. "Race and Ethnicity and the Controversy over the US Census." *Current Sociology* 48 (3):87–110.

Bauman, Zygmunt, Didier Bigo, Paulo Esteves, Elspeth Guild, Vivienne Jabri, David Lyon, and R. B. J. Walker. 2014. "After Snowden: Rethinking the Impact of Surveillance." *International Political Sociology* 8 (2):121–144.

Bigo, Didier. 2011. "Pierre Bourdieu and International Relations: Power of Practices, Practices of Power." *International Political Sociology* 5 (3):225–258.

Bigo, Didier. 2013. "The Transnational Field of Computerised Exchange of Information in Police Matters and its European Guilds." In *Transnational Power Elites: The Social and Global Structuration of the EU*, edited by Niilo Kauppi and Mikael Rask Madsen, 155–182. London: Routledge.

Bigo, Didier, and R.B.J. Walker. 2007. "Political Sociology and the Problem of the International." *Millennium: Journal of International Studies* 35:725–739.

Bonditti, Philippe, Didier Bigo, and Frederic Gros, eds. 2017. *Foucault and the Modern International*. New York: Palgrave MacMillan.

Bourdieu, Pierre. 1973. "L'opinion publique n'existe pas." *Les temps modernes* (318):1292–1309.

Bourdieu, Pierre. 1988. "Social Space and Symbolic Power." *Sociological Theory* 7 (1):14–25.

Bourdieu, Pierre. 1993. *Language and Symbolic Power*. Cambridge, MA: Harvard University Press.

Bourdieu, Pierre, and Loïc J. D. Wacquant. 1992. *An Invitation to Reflexive Sociology*. Chicago, IL: University of Chicago Press.

Butler, Judith. 1997. *Excitable Speech: A Politics of the Performative*. London: Routledge.

Desrosières, Alain. 1998. *The Politics of Large Numbers: A History of Statistical Reasoning*. Translated by Camille Naish. Cambridge, MA: Harvard University Press.

Dezalay, Yves, and Bryant G. Garth. 1996. *Dealing in Virtue: International Commercial Arbitration and the Construction of a Transnational Legal Order*. Chicago, IL: University of Chicago Press.

Foucault, Michel. 2007. *Security, Territory, Population*. Translated by Graham Burchell. Edited by Arnold Davidson, *Lectures at the Collège de France, 1977–78*. Basingstoke: Palgrave Macmillan.

Galloway, Alexander R. 2006. *Protocol: How Control Exists After Decentralization*. Cambridge, MA: MIT Press.

Gandy, Oscar, and Lokman Tsui. 2018. "On Personal Data Protection, Privacy and Surveillance." *Communication & Society* 43:1–34.

Hacking, Ian. 1990. *The Taming of Chance*. New York: Cambridge University Press.

Hacking, Ian. 2015. "Biopower and the Avalanche of Printed Numbers." In *Biopower: Foucault and Beyond*, edited by Vernon W. Cisney and Nicolae Morar, 65–80. Chicago, IL: University of Chicago Press. Original edition, 1983.

Haraway, Donna Jeanne. 1991. *Simians, Cyborgs and Women: The Reinvention of Nature*. London: Free Association.

Isin, Engin, and Evelyn Ruppert. 2015. *Being Digital Citizens*. London: Rowman & Littlefield.

Kertzer, David I., and Dominique Arel. 2002. "Censuses, Identity Formation, and the Struggle for Political Power." In *Census and Identity: The Politics of Race, Ethnicity, and Language in National Censuses*, edited by David I. Kertzer and Dominique Arel, 1–42. Cambridge/New York: Cambridge University Press.

Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: SAGE.

Lake, Robert. 2017. "Big Data, Urban Governance, and the Ontological Politics of Hyper-Individualism." *Big Data & Society* Jan–June:110.

Lyon, David. 2014. "Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique." *Big Data & Society* 1:1–13.

Madsen, Mikael Rask. 2011. "Reflexivity and the Construction of the International Object: The Case of Human Rights." *International Political Sociology* 5 (3):259–275.

Madsen, Mikael Rask. 2014. "The International Judiciary as Transnational Power Elite." *International Political Sociology* 8 (3):332–334.

Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution that will Transform How we Live, Work and Think*. London: John Murray.

Monbiot, G. 2017. "Big data's power is terrifying. That could be good news for democracy." The Guardian, 6 March 2017. Accessed 2 August 2017. www.theguardian.com/commentisfree/2017/mar/06/big-data-cambridge-analytica-democracy

Nobles, Melissa. 2000. *Shades of Citizenship: Race and the Census in Modern Politics*. Stanford, CA: Stanford University Press.

Osborne, Thomas, and Nikolas Rose. 1999. "Do the Social Sciences Create Phenomena? The Example of Public Opinion Research." *British Journal of Sociology* 50 (3):367–396.

Porter, Theodore M. 1986. *The rise of statistical thinking, 1820–1900*. Princeton, NJ: Princeton University Press.

Ruppert, Evelyn, Engin Isin, and Didier Bigo. 2017. "Data Politics." *Big Data & Society* 4 (1):1–8.

Scheel, Stephan, Baki Cakici, Francisca Grommé, Evelyn Ruppert, Ville Takala, and Funda Ustek-Spilda. 2016. *Transcending Methodological Nationalism through Transversal Methods? On the Stakes and Challenges of Collaboration*. ARITHMUS Working Paper No. 1. Goldsmiths University of London. Avail. at http://bit.ly/2lqR1aM.

Socialising Big Data Project. 2015. *A Social Framework for Big Data*. CRESC. University of Manchester and the Open University. Avail. at http://bit.ly/28Ye3kI.

Thrift, Nigel. 2005. *Knowing Capitalism*. London: SAGE.

Toxen, B. 2014. "The NSA and Snowden: Securing the All-Seeing Eye." *Communications of the ACM* 57 (5):44–51. DOI: 10.1145/2594502.

PART I

Conditions of possibility of data politics

2

KNOWLEDGE INFRASTRUCTURES UNDER SIEGE

Climate data as memory, truce, and target

Paul N. Edwards

Introduction

Data politics are not really about data per se. Instead, the ultimate stakes of how data are collected, stored, shared, altered, and destroyed lie in the quality of the knowledge they help to produce. The value of knowledge depends on trust in the infrastructures that create it, which themselves depend on trusted data. "Trust" and "truth" are closely related English words, and both are strongly connected to the "truce" of my title. Indeed, etymologically "true" and "truce" are in fact *the same* word.[1]

This chapter argues that both trust in climate knowledge infrastructures and the truth they deliver descend from the truces necessary to share and maintain climate data—all of them heavily contested in the contemporary United States. The partial success of climate change denialism stems, in large part, from the recent arrival of what I call "hypertransparency": open data, open code, commodity software tools, and alternative publication venues have quite suddenly upended truces painstakingly built over multiple centuries. Not only climate knowledge, but virtually all scientific knowledge of public concern is affected by this transformation.

My title derives from Nelson and Winter's theory of organizational routines in *An Evolutionary Theory of Economic Change* (1982). In large organizations, Nelson and Winter posited, routines serve three critical purposes. First, they hold the organization's knowledge, storing organizational *memory* even as individual employees come and go. Argote (1999) later argued that automating routines transfers organizational knowledge to machines, a line of reasoning readily extended to many kinds of algorithms. Second, routines and technologies constitute a de facto *truce* among individuals and organizational elements whose goals and interests conflict. Routines, machines, and algorithms may incorporate workarounds and exception-handling

procedures, and organizational culture can evolve to tolerate failures, delays, and protests (up to a point). Finally, routines, machines, and algorithms function as *targets*, in two senses. First, they embody the organization's aims and goals. Second, they serve as patterns or templates for managing new systems.

Nelson and Winter's paradigm cases were factories and corporations, but their theory fits most organizational contexts. It can be extended to encompass routines that cross organizational boundaries, as is typical in the case of infrastructures (Edwards et al. 2007). In the first part of this chapter, I show how the routines, machines, and algorithms of climate data systems fit Nelson and Winter's paradigm of memory, truce, and target. Against this background, I then engage a much less benign sense of the word "target"—namely, "an object aimed at in shooting"—in the context of aggressive assaults on environmental knowledge infrastructures in the contemporary United States. These attacks seek to dismantle the very sources of climate knowledge by defunding crucial instruments, satellites, and data analysis programs.

"Long data" and environmental knowledge

Some kinds of environmental knowledge, such as the monitoring of weather, water quality, air pollution, or seismic activity, concern what is happening in the present, often with a view to short-term prediction (hours to days). Many such systems collect "big data," e.g. from satellites or large sensor networks. As environmental data make their way into archives, they become "*long* data" (Arbesman 2013). Long data enable scientists to track and understand environmental change.

Memory: environmental data and data models

Enduring, carefully curated collections of long data—scientific memory—are therefore among the most valuable resources of modern environmental science. Such collections require continual maintenance. Over time, all kinds of things change about how data are collected. Instruments and sensors get out of calibration, or are replaced with newer models with different characteristics. New human observers, new standards, new organizations, and new methods replace predecessors. National boundaries and political systems also change, sometimes with effects on data-collecting practices (Edwards 2010; Gitelman 2013).

Such changes nearly always affect the relationship of today's data to those recorded in the past. If a new digital thermometer (say) systematically reads 0.2°C higher than a previous model, this is because the instruments differ, not because the air temperature has actually changed. To keep the long-term record consistent, scientists must adjust data to take account of these shifts. In the simple case just mentioned, when graphing temperature change over time, they might add 0.2°C to all readings taken by the older instrument. Far from falsifying data, this practice actually "truthifies" the long-term record. Adjusting data was once a manual routine; today it is usually handled by computerized algorithms. I often call such algorithms "data models," because they encode models of how different instruments or data sources relate to each other.

In addition, many environmental sensing instruments, such as those flown on satellites, produce data that *require* interpretation by algorithms before they can be used for most purposes. Since data analysis algorithms change—often but not always for the better—data interpreted with older algorithms must be continually re-interpreted using newer ones. For example, many major satellite data sets are "versioned," or re-issued after reprocessing with revised algorithms. At this writing, the 35-year record of tropospheric temperatures from satellite-mounted microwave sounding units is on version 6, with multiple minor releases between major ones (Spencer, Christy, and Braswell 2017).

Routines for homogenizing data—whether enacted by people, machines, or algorithms—embody scientific organizations' memory of how their data were previously created, just as Nelson and Winter argued. Perhaps counterintuitively, this ongoing adjustment—and revisions of the techniques used to make those adjustments—proves critical to maintaining and even improving the integrity of the long-term record. Different versions of data sets often, though not always, converge on similar results as analysis techniques improve. Further improvement—i.e., further revision, potentially reversing previous convergence—is almost always possible, and no data set is ever entirely definitive. In *A Vast Machine* (2010), I called this phenomenon "shimmering data."

Truce: keeping things running

Nelson and Winter's idea of routines as truces among conflicting groups and goals appears very clearly in the record of many environmental data systems. An extreme example of one such truce: even during the Cuban Missile Crisis in 1962, the weather services of Cuba, the USSR, and the United States continued to exchange weather data (Schatz 1978). With only a few exceptions, from the late 19th century to the present only actual war interrupted the routine flow of meteorological data among national weather services—even those of pariah states such as North Korea.

When a scientific field depends on data from many sources, such as the weather services of the world's 190-plus nations, conflicts among the various contributors are almost inevitable. In meteorology, such conflicts have occurred over observing hours, temperature and pressure scales, instrument placement, instrument housings, methods of calculating averages, reporting deadlines, and many other details of exactly when, where, and how observational data are taken and recorded. To share data in any meaningful way, these differences must be reconciled (Bowker 2000; 2005). Before computers, this was done primarily by standard-setting and laborious *post hoc* calculation; after computers, algorithms took over the calculations.

The dramatic advance of the environmental and geosciences resulted in large part from routines-as-truces surrounding data. In meteorology, merchant ships from many nations began keeping standardized weather logs in the 1850s; these logs represented a truce among nations that used different temperature scales and other norms. The International Meteorological Organization, formed in the 1870s,

promoted common standards and data sharing (Daniel 1973). The World Data Centers established during the 1957–58 International Geophysical Year, at the height of the Cold War, today remain critical institutions supporting virtually all of the geosciences: soil, solid Earth, climate, oceanography, and others (Aronova 2017; International Council of Scientific Unions 2016).

Examples are easily multiplied.

Target: routines and technologies as patterns for new systems

By representing the new and unfamiliar in terms of something old and well understood, the page, file folder, and wastebasket icons of modern computers helped enable the transition from paper-based office information systems. Similarly, Nelson and Winter argued that when organizations confront new needs, they use their existing routines as "targets" for the design of new ones. In just this way, existing data handling routines serve as targets when new data sources become available.

For example, satellite radiometers can measure atmospheric temperature, but they do so in a very different way from instruments that work by direct contact with the atmosphere, such as the thermometers carried on weather balloons. Essentially, they measure huge volumes of air, rather than taking readings at points along an ascending line. When satellite radiometry first became available in the 1970s, meteorologists developed algorithms to make satellite data look like weather balloon data in order to incorporate them into existing weather forecast models. Only much later were techniques developed to ingest satellite data in a form more appropriate to what they actually measure.

By now the overall point should be clear. Scientific memory requires truces: acceptance of common standards, suspending conflicts and disagreements to get on with routine data sharing, putting aside political and ideological differences to work together. Such truces—and the mature technological systems that embody them—endow knowledge infrastructures with considerable stability and inertia (Edwards et al. 2007). Usually (but not always) this inertia prevents sudden, dramatic changes in how knowledge is created, disseminated, and understood. Like other organizational routines, they bear history and memory within them, maintaining coherence and providing templates for new data systems as they arise. Only with such truces can a knowledge commons emerge, beyond both organizational and national borders (Edwards 2013).

The glass laboratory

My argument in this section is that in the early 21st century, certain truces became targets—not in Nelson and Winter's benign sense, but in the more destructive sense of "things to shoot at."

To understand the siege I will discuss in Part III, this section first offers some general background on the evolving character of transparency in science. Next, I

briefly discuss three climate controversies of the early 21st century. These include the "hockey stick" graph of global temperature since AD 1000, the citizen-science project surfacestations.org, and the Berkeley Earth Surface Temperature (BEST) project. Together, these examples illustrate the emergence of *hypertransparency* in climate science: a "glass laboratory" in which virtually every aspect of scientific work—data, algorithms, software, email—takes place in a vastly expanded public arena. In this condition, the truces that made modern climate science possible come under severe, sometimes disabling scrutiny, with both negative and potentially positive consequences for public knowledge.

Transparency as norm: from open methods to open data

Merton (1973) famously argued that norms of disinterestedness and communalism guide science: scientific findings belong to the entire community, rather than to individual investigators. This principle manifests in publication, a word that means, quite literally, "making public." In the tradition of "virtual witnessing" established by the Royal Society of London in the 17th century—eloquently elaborated by Shapin and Schaffer in *Leviathan and the Airpump* (1985)—modern science "makes public" not only its results, but also the methods by which they were obtained. Scientific articles enrol the reader as a "virtual witness" to an openly and fully described process.

As Shapin and Schaffer also showed, however, the "public" in question was originally a very limited, elite group. In the case of the Royal Society, only "gentlemen" could participate, whether as experimenters or as trusted witnesses. By the second half of the 20th century, the practice of virtual witnessing had evolved to include what we call peer review, the gateway to publication. In this system, other scientists doing closely related work serve as initial referees (virtual witnesses), while the assumed Mertonian norms of disinterestedness and scepticism replace those of the gentlemanly "modest witness." Once approved and published, any member of "the public" could (in principle) conduct a new trial of the results, using the methods section as a recipe.

These were never, of course, more than ideals. One need no longer be a gentleman to participate, but not just anyone can do so; one must still be a practicing scientist. Most people—even most scientists outside the focal discipline of any given publication—lack the extensive training, specialized vocabularies, and access to laboratories and other resources required even to understand, much less to reproduce, any given result. On top of that, the descriptions in methods sections almost never contain enough information to permit exact replication (Collins 1985; Collins and Pinch 1993). Peering through the tiny window of a scientific article, one can't see everything.

With respect to data, until recently most scientific publications fell far short of full transparency. In most fields, so-called "raw" data were never revealed at all. Instead, scientists published graphs, tables, or charts that synthesized large bodies of data. Practicalities partly explain this: raw data could be voluminous and expensive

to publish on paper, yet even expert readers were unlikely even to look at them. Peer reviewers had the right to review data, but in practice they rarely did so. Together, these cultural practices made raw data effectively the intellectual and physical property of their creator (Bowker 2000, 646).

The environmental and geosciences exhibited a different pattern, with many data published and freely shared. Yet even in those fields, a sense of personal ownership is not uncommon. When I first started researching the history of climate science in the mid-1990s, I was regularly directed to "data guys" who specialized in one kind of instrument, region, or record. The moniker "data guys" reflected these individuals' deep knowledge of quirks and limitations, down to specific events in the life of particular surface stations or instruments. The phrase also implied a sense of ownership: only these "data guys" really "knew" the data in question.

In summary, *limited* transparency for *limited* publics has always played a major role in modern science. *These limits on visibility and participation were critical, widely accepted truces*. They were key to remarkable successes in producing reliable knowledge. Yet at different times and in different contexts, those same limits have led to both trust and distrust, broad benefits and broad harms, social uplift and damaging elitism (Jasanoff 2005; Lövbrand, Pielke, and Beck 2010; Miller 2007). The tacit social agreements making up those truces have been thoroughly explored by science and technology studies.

Less widely remarked have been the equally critical limits on participation imposed by access to technologies and skills—limits that have diminished dramatically in recent years. Breathtaking Moore's-law declines in the cost of storing digital data and the rise of high-speed computer networks have removed most practical obstacles to publishing or sharing even large data sets and scientific software. By the early 2010s, these factors combined with a perceived "replication crisis" in science to create pressures for scientists to "publish everything": data, software, spreadsheets, experimental protocols, etc. (Baker 2015). Some journals and many science funders began to require would-be authors to deposit both data sets and computer models or scripts along with articles describing their findings (Stodden, Seiler, and Ma 2018). Across the same period, scientific tools for data analysis became cheaper, simpler, and much more widely available. Examples include statistical analysis software, spreadsheets, visualization tools, and scripting languages. Millions of people acquired skills in statistics and in coding, both once arcane arts.

Most of these phenomena extended, of course, far beyond science. Open source code and open access to data, software, and publications were part of a generalized clamour for greater transparency in government and even in corporate life (Goldstein, Dyson, and Nemani 2013; Weinberger 2012). Instead of mere windows, it suddenly seemed possible to replace every laboratory (or government) wall with glass, and to move this glass laboratory (metaphorically) from a quiet campus to the centre of a great city. Now "the public" could include anyone with access to a computer and the Internet. Credentialed visitors, informed spectators, random passers-by, and snarling vandals could all play the role of virtual witness. Promoters of transparency typically invoked the values of democracy, participatory process, and

institutionalized scepticism ("trust but verify")—but vastly increased transparency has proven an equal if not greater boon to conspiracy theorists, muckraking journalists, and populist politicians.

Truces as targets: three climate data controversies of the early 21st century

As ever more virtual witnesses crowded around to watch climate science in action, some chose to join in, to test and sometimes to challenge methods and data. In this section, I show how the glass laboratory revealed some of the many truces involved in "making data global," as I put it in *A Vast Machine* (2010, chapter 10)—and how the rhetoric of transparency led to outcomes that presage today's concerted assault on the climate knowledge infrastructure. Here I have space to offer only sketches of each controversy, pointing the reader to other sources for more complete discussions.

Climate Audit

If you are a corporation, you need to get your books audited regularly. Internal audits are fine for some purposes, but they create an opening for corruption, because an auditor whose salary you pay has a strong incentive to tell you what you want to hear. That auditor might be tempted to alter, conceal, or delete data that doesn't look good for your bottom line, or you. Therefore, to show shareholders and regulators that you're telling the truth, you need an *independent* audit. Further, it's the job of auditors to check that results are correctly derived from data.

This was the idea behind "Climate Audit," a blog started in 2004 by retired Canadian mining engineer Steve McIntyre. Peer review is fine, he might say, but since this week's peer reviewer is next week's peer reviewee, scientists have a certain incentive to approve each other's work: peer review can become "pal review." Even if scientists aren't doing each other personal favours, there remains the potential for "groupthink" (Janis 1982): what Kuhn (1962) might have characterized as collective commitment to a paradigm, causing an entire field to ignore anomalous data or results. Peer reviewers influenced by groupthink might resist publication of results that violate the paradigm. Funding agencies might also be victims of groupthink, in a positive feedback loop that starves out research aimed at disconfirming a widely accepted theory. Since climate change is a major public policy issue, McIntyre might argue, important results need review by some neutral party without financial or personal stakes in the outcome.

The first of McIntyre's "audits" involved the "hockey stick" graph of northern hemisphere average temperatures (see Multiproxy line in Figure 2.1, called a hockey stick because the long "handle" of relatively little variation ends in a "blade" of steeply rising average temperatures in the 20th century). This graph, first published in Mann, Bradley, and Hughes (1998) for the period 1400–1980 AD, was subsequently extended to encompass the period since 1000 AD (Mann,

Bradley, and Hughes 1999, hereafter MBH99). MBH99 used statistical techniques to combine some 112 proxy data sets, such as corals, tree rings, and ice cores, and historical thermometer records. (Proxies are things that vary consistently with temperature, so that they can be used as surrogate thermometers.) Their results showed recent temperatures as more than 0.5°C higher than at any time in the last 1,000 years (Houghton 2001; Mann and Jones 2003).

Working with economist Ross McKitrick and an eventual army of interested blog readers, McIntyre requested and received Mann's raw data, along with the "weights" applied to different proxies in his analysis. Without claiming expertise in paleoclimate studies, McIntyre and McKitrick argued that they were still competent to evaluate the MBH studies since "the same algebraic and statistical methods [used by MBH] are commonly used in economics, business and elsewhere in the social sciences." Interrogating the MBH data and analysis, they claimed to find significant problems in both, including "collation errors, unjustifiable truncation or extrapolation of source data, obsolete data, geographical location errors, incorrect calculation of principal components and other quality control defects." McIntyre and McKitrick concluded that the sharp upward curve seen in the MBH98 northern hemisphere temperature graph was "primarily an artefact of poor data handling, obsolete data, and incorrect calculation of principal components" (McIntyre and McKitrick 2003). Their own analysis seemed to show that average northern hemisphere temperatures were higher in the 1400s than at any other time before 1980, when the analysis ended.

Thus began a complicated and acrimonious dispute that continues, in some quarters, to this day. Mann admitted to certain errors, but disputed most others and held that correcting those errors made little difference to the final outcome (Mann, Bradley, and Hughes 2004). McIntyre and McKitrick amped up their critique of the MBH methods, gradually shifting from a polite, participatory mode into a much more adversarial discourse that stopped just short of accusing MBH of outright fraud (McIntyre and McKitrick 2005).

These disputes eventually led to hearings before the US House of Representatives Subcommittee on Oversight and Investigation, as well as reviews by the US National Science Foundation, the American Association for the Advancement of Science, the National Academy of Sciences, and the US National Research Council, as well as others. In 2006, a lengthy NRC report concluded that "uncertainties of the published reconstructions have been underestimated," expressing only lukewarm confidence in proxy-based reconstructions of the period 1000–1600 AD. However, it also noted that multiple lines of evidence did support the conclusion that the warmth of the late 20th century "in many cases appear[s] to be unprecedented during at least the last 2,000 years." Despite large uncertainties, multiple independent proxy reconstructions also tended to support the MBH conclusions, according to the report (Committee on Surface Temperature Reconstructions for the Last 2000 Years, Board on Atmospheric Sciences and Climate, and National Research Council 2006). As a direct result of this "audit," the unfortunate Michael

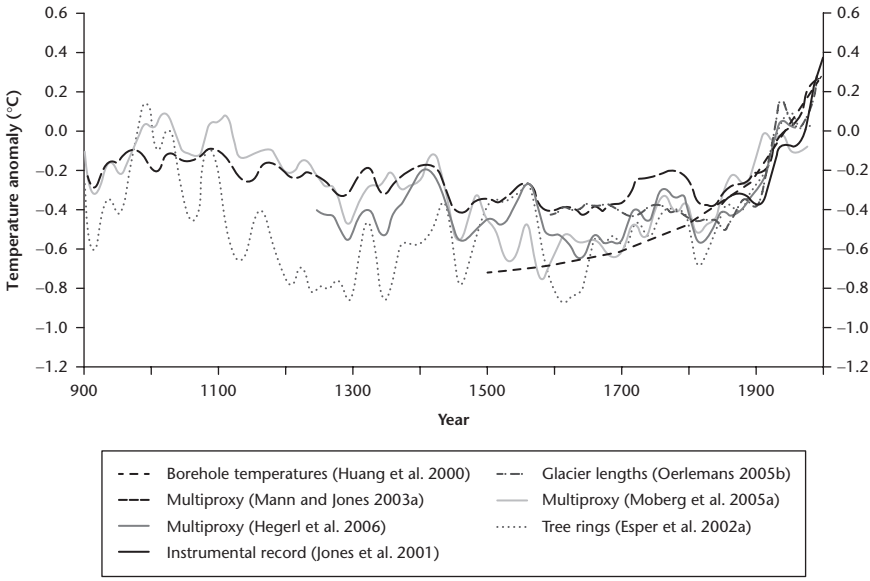


FIGURE 2.1 Northern Hemisphere Average Temperatures[2]

Mann became a *bête noire* of climate change deniers, subjected to multiple lawsuits, investigations, harassment, and death threats (Mann 2012).

Meanwhile, McIntyre's blog gained tens of thousands of followers, becoming a kind of clearinghouse for both genuinely sceptical but open-minded readers and ideologically-driven climate change deniers. In November 2009, Climate Audit was among the first websites to publish the hacked "Climategate" emails. Not coincidentally, earlier that year McIntyre and various Climate Audit readers had filed 58 Freedom of Information requests to the Climatic Research Unit at the University of East Anglia, from which the emails were stolen (Pearce 2010).

Climate Audit made an explicit goal of disrupting key truces in the climate knowledge infrastructure as it then existed. Early on, McIntyre and McKitrick excoriated the existing peer review system and proposed the audit as a new, higher standard:

> We are . . . struck by *the extremely limited extent of due diligence involved in peer review* as carried out by paleoclimate journals, as compared with the level of due diligence involved in auditing financial statements or carrying out a feasibility study in mineral development. For example, "*peer review" in even the most eminent paleoclimate publications, as presently practiced, does not typically involve any examination of data, replication of calculations or ensuring that data and computational procedures are archived.* We are not suggesting peer reviewers should be auditors. Referees are not compensated for their efforts and journals would not be able to get unpaid peer reviewers to carry out thorough audits.
>
> *(McIntyre and McKitrick 2005, 94, emphasis added)*

If peer reviewers need not be auditors, then who? McIntyre and McKitrick's forays into publishing in climate science suggested one approach, i.e. publishing their critique in a traditional venue, itself subject to peer review. But the example of McIntyre's blog—from its title to its contents and hard-hitting tone—suggested another: that a random crowd of deeply sceptical, variably skilled, often anonymous interpreters might do the trick. The fact that the thorny problem of groupthink, mentioned earlier, might prove even harder to avoid in the case of a public blog did not dissuade him from pursuing that course.

Surfacestations.org

Without necessarily using the "audit" vocabulary, numerous similar projects have ensued since about 2005. One is surfacestations.org, initiated in 2007 by television meteorologist Anthony Watts, who runs the extremely popular sceptic blog WattsUpWithThat (WUWT). From 2007–2017, according to Watts, WUWT racked up more than 316 million page views and 2 million comments on some 16,500 posted "stories."[3] Watts believed, and apparently still believes, that at least part of the warming trend since the 1970s is due to an upward bias in the surface temperature record due to poorly sited weather stations. To test this hypothesis, Watts used his blog to enlist volunteers in an audit of all stations in the US Historical Climatology Network (USHCN), a specialized subset of weather stations selected for "their spatial coverage, record length, data completeness, and historical stability."[4]

The ideal site (location) for weather instruments is a flat area surrounded only by grass or low vegetation, far from large bodies of water, buildings, and anything else that might render readings unrepresentative of the surrounding area. An imperfect site can subject thermometers to various forms of bias, such as artificial heating from asphalt parking lots. The National Climatic Data Center's *Climate Reference Network (CRN) Site Information Handbook* defines five classes of sites: classes 1 and 2 are least susceptible to bias, whereas classes 3, 4, and 5 exhibit biases above 1°C, 2°C, and 5°C respectively. A class 5 site might, for example, have a thermometer located near a building's heating vent or air conditioner exhaust.

Over 650 volunteers signed on. Watts instructed them to visit stations with cameras and the *Site Information Handbook* in hand. They wrote up their results according to the *Handbook's* rating scheme and provide detailed photographic and written evidence. They eventually surveyed nearly every station in the USHCN.

The surfacestations.org survey seemed to validate Watts's suspicions. He released a "midterm report" entitled "Is the U.S. Temperature Record Reliable? How do we know global warming is a problem if we can't trust the U.S. temperature record?" published by the denialist thinktank Heartland Institute (2009). It touted the "inescapable" conclusion that "the US temperature record is unreliable" and claimed that it "reports a false warming trend."

Data from the citizen-science survey, including photographs, were posted on the surfacestations.org website. Quantified, the results portrayed a network with serious problems. A mere 11 percent of stations fell into the minimally-biased classes 1 and 2, while a whopping 69 percent fell into classes 4 and 5—heavily biased warm.

The story then took an unusual turn. Researchers at the US National Climatic Data Center (NCDC), led by Matthew Menne, downloaded and analysed the surfacestations.org data. For a key subset of stations, they compared the surfacestations.org ratings with independent ratings of the same stations by the National Weather Service. For the most part, these ratings agreed: Watts's volunteers had done an excellent job. Of the 525 stations they evaluated, only 71 fell into the "good" classifications (ratings 1 or 2).

Oddly, however, when Menne's group compared the temperature trends calculated from the "good" and "poor" stations, they found that *unadjusted* temperatures for the "poor" stations *actually exhibited a slightly cool bias* relative to the "good" stations. This counterintuitive result turned out to be due to specific characteristics of the temperature sensors. Most of the "poor" stations used Maximum/Minimum Temperature System (MMTS) sensors, a relatively new type installed since the mid-1980s. These electronic sensors are attached by cables to an indoor readout device. Limits on the maximum cable length required that many be installed much closer to buildings and parking lots than the optimal siting guidelines allowed. At most of the "good" stations, by contrast, thermometers were an older liquid-in-glass type installed in Cotton Region Shelters located further from distorting influences. However, *the MMTS sensors register lower maximum temperatures than the CRS type*, accounting for the cool bias at the apparently "poor" station sites. The NCDC researchers then showed that the algorithms used to adjust data across the entire network already produced very close agreement between "good" and "poor" stations—though even after adjustment, a slight cool bias in the "poor" stations remained (Menne, Williams, and Palecki 2010).

In this example, two kinds of truce became targets. First, the raw data themselves came under scrutiny by citizen scientists who challenged Weather Service data collection systems, demonstrating justifiable concerns about data quality. Given WUWT's relentlessly denialist posture, most of these volunteers probably hoped to discredit Weather Service data. Second, Weather Service data modelling—designed to correct for missing data and inevitable biases such as those introduced by replacing CRS sensors with MMTS—came under fire. Watts's "midterm report" noted that data adjustment routines included replacing missing station data with data from other, nearby stations (an old practice in weather data analysis), and claimed that "adjustments applied to 'homogenize' the data . . . impart an even larger false warming trend to the data" (Watts 2009, 13).

This episode ended with acceptance and integration of the citizen-science data into normal scientific routines. Watts later co-authored a paper with John Christy, Roger Pielke Sr., and other sceptical scientists. After peer review, Watts's original claims of enormous bias in the USHCN were drastically tempered:

> Temperature trend estimates vary according to site classification, with poor siting leading to an overestimate of minimum temperature trends and an underestimate of maximum temperature trends. . . . The opposite-signed differences of maximum and minimum temperature trends are similar in magnitude, so that *the overall mean temperature trends are nearly*

> *identical across site classifications*. Homogeneity adjustments tend to reduce trend differences, but statistically significant differences remain for all but average temperature trends.
>
> *(Fall et al. 2011, D14120, emphasis added)*

Within the scientific community, then, this audit ended by essentially corroborating the NCDC's data, its data collection routines, and its data models, while pointing to room for small improvements in both.

The mother of all audits: Berkeley Earth

Among the most intriguing and thorough "audits" of climate data is the Berkeley Earth Surface Temperature project (BEST, also known simply as Berkeley Earth), initiated in 2010 by UC Berkeley physics professor Richard Muller. A self-proclaimed "agnostic" on global warming, Muller had been an early critic of the MBH "hockey stick" graph, finding merit in the arguments of McIntyre, McKitrick, and Watts. Following Watts's surfacestations.org study and the "Climategate" email controversy, Muller decided to revisit the global land surface temperature record. He sought and received funding for this effort from several sources, including the US Dept. of Energy, Bill Gates, and the Bowes, Folger, and Getty foundations. The largest single contribution ($150,000) came from the ultraconservative Charles Koch Foundation, which doubtless hoped that Muller would put the lie to global warming science. Convinced of Muller's neutrality, Anthony Watts contributed data from surfacestations.org to the project. On his blog, Watts wrote:

> *I'm prepared to accept whatever result they produce, even if it proves my premise wrong . . .* My gut feeling? The possibility that we may get the elusive "grand unified temperature" for the planet is higher than ever before. Let's give it a chance.[5]

BEST collected data from some 39,000 weather stations—over five times as many as any previous study. These data included many records previously considered too brief, too intermittent, too biased, or too poorly standardized for use in climate studies. BEST developed its own quality control algorithms to eliminate impossible readings, duplicate records, and other problems. It also developed new algorithms for adjusting and homogenizing data. The project made a special point of maximum transparency, publishing all of its data, algorithms, and code online at berkeleyearth.org.

Like Watts, Muller released his preliminary findings before peer review—at a March 2011 session of the US House Committee on Science, Space and Technology. The trend seen in the BEST analysis, he told the representatives, "is very similar to that previously reported by the other [climate data] groups" (Roosevelt 2011). In fact, the BEST analysis differs hardly at all from those of the three major climate data stewards: NASA's Goddard Institute for Space Studies,

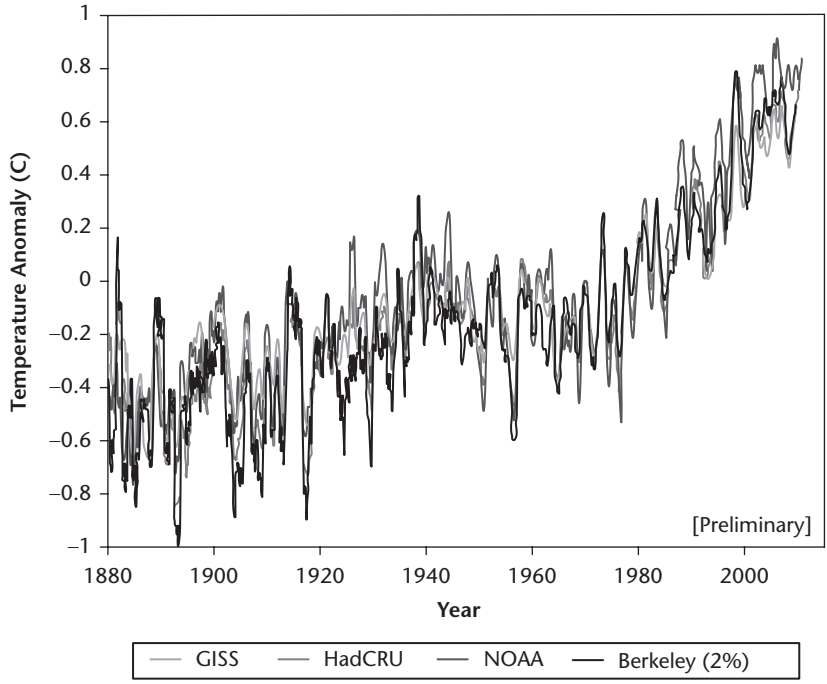


FIGURE 2.2 BEST Annual Land-Surface Temperature Analysis[6]

NOAA, and the UK's Hadley Centre and Climatic Research Unit (Figure 2.2). BEST noted that the urban heat island effect (related to Watt's critique of surface station bias) is "large and real," but concluded that it "has not had a significant impact on global temperatures since 1950," primarily because urban areas take up only a tiny fraction of Earth's total land area (Wickham et al. 2013: 6).

Far from accepting Muller's results, as he had promised the previous month, Watts now criticized Muller for presenting them prior to peer review—even though he had done the same thing with his own "midterm report." Yet when it appeared in late 2012, the final peer-reviewed study differed only in minor details from what Muller showed to Congress (Muller et al. 2013). Watts now cites Muller's pre-review release to justify doing the same.

New fronts in the siege

Taken together, these three episodes demonstrated how digital technologies, combined with open data policies, have empowered serious challenges to knowledge infrastructures controlled by government and academic institutions. "Climate audits" were only part of an older, much larger industry-led attack on environmental science with regulatory implications. Deliberate efforts to undermine scientific conclusions about climate change date to the early 1990s and even

before; in other areas, similar efforts can be traced to the dawn of environmental regulation in the 1970s (Jasanoff 1990). This picture includes major lobbying efforts and public relations campaigns funded by fossil fuel interests and wealthy donors, often via conservative think tanks such as the Heartland Institute and the George C. Marshall Institute (Hoggan and Littlemore 2009; Oreskes and Conway 2010).

The so-called "Nongovernmental International Panel on Climate Change"—sponsored by the Heartland Institute—makes the stakes quite clear: the NIPCC is

> an international panel of *nongovernment scientists and scholars* who have come together to understand the causes and consequences of climate change. Because we are not predisposed to believe climate change is caused by human greenhouse gas emissions, we are able to look at evidence the Intergovernmental Panel on Climate Change (IPCC) ignores. *Because we do not work for any governments, we are not biased toward the assumption that greater government activity is necessary.*
>
> *(emphasis added)*[7]

The obvious, but false implication is that IPCC members are "government scientists" whose hidden goal is to promote a restrictive regulatory agenda, which could be damaging to the fossil fuel industry, as well as others. In reality the 2,500-plus scientists who work on each IPCC report participate as unpaid volunteers. Their salaries and research support come from their home institutions, which may be academic, industrial, or government laboratories. Government funding mechanisms and IPCC authorship protocols are designed to *prevent* political (including government) interference with the scientific process. Representatives of national governments do participate directly in final editing and approval of IPCC summaries but this review comes only after the scientific conclusions have been finalized. Government influence is widely believed to have pushed the reports toward *more conservative* conclusions (Edwards and Schneider 2001).

Politically motivated attacks on the conclusions of climate science take many other forms as well. The most common strategies include repeating false claims long after they have been disproven, maligning and harassing individual scientists, altering official reports (as was done on multiple occasions during the George W. Bush administration), and attempting to muzzle scientists employed by government agencies (Bowen 2008; Mann 2012; Office of Inspector General, National Aeronautics and Space Administration 2008). Lacking space to discuss all of this, my final section focuses specifically on current attempts to undermine the climate knowledge infrastructure.

In 2017, following the inauguration of President Donald Trump, the siege progressed to a new stage, with a new series of targets. This new approach goes far beyond the questioning of knowledge *outputs* (conclusions, estimates of uncertainty, data analyses, etc.). Instead, it seeks to delete important *inputs* to the knowledge-making process, such as key instruments required to monitor environmental change, and to remove climate change from the missions of

agencies charged with environmental monitoring. In some cases, it has forcibly restructured review processes, altering who has standing to act as a peer in reviewing environmental policies based on science. If successful, this siege will certify poorly qualified people as expert peers while decertifying highly qualified scientists; introduce significant discontinuities and/or distortions in the long-term data record; and diminish the quality and quantity of public knowledge about the state of the environment.

Instruments: eliminating satellites and their data

The National Aeronautics and Space Administration (NASA) currently holds responsibility for designing and launching satellites carrying environmental sensing instruments. In April 2017, in its first budget proposal, the incoming Trump administration proposed cuts of about US$167 million to NASA's budget for earth science. Virtually all of that would come from eliminating five specific satellite missions directly related to climate science:

- RBI (Radiation Budget Instrument), providing data to establish the amount of incoming and outgoing radiation, crucial to both long-range weather forecasts and climate sensitivity studies
- PACE (Plankton, Aerosol, Clouds and Ocean Ecosystem), providing data relevant to ocean health, cloud formation, and the carbon cycle
- OCO-3 (Orbiting Carbon Observatory), monitoring the location of carbon sources and sinks
- CLARREO (Climate Absolute Radiance and Refractivity Observatory) Pathfinder, a solar spectrometer
- DSCOVR (Deep Space Climate Observatory) provides radiation monitoring as well as high-resolution images of the Earth and its moon.

The last of these, nicknamed "Goresat," was funded and built under President Clinton, but mothballed before launch by President G.W. Bush. Revived under President Obama, it was finally launched in 2015. The same satellite also carries two NOAA space weather instruments. The Trump budget would have eliminated funding for data analysis of *only* the two DSCOVR instruments run by NASA, i.e., those proposed by Gore (a pariah for the far right). Savings from the DSCOVR cut: a mere $3.2 million. As one policy analyst dryly put it, "budget savings do not appear to be the driving force for proposing these terminations. Instead, it apparently is based on Trump Administration scepticism about climate change" (Smith 2017, 5).

Under the complex US budget process, the proposed cuts were approved by the House of Representatives, but specifically and entirely restored by the Senate before the FY2018 budget was signed into law. NASA terminated the RBI program on its own initiative in January 2018, citing cost overruns and the existence of other, similar instrument. The administration's FY2019 budget once again proposed cutting the four

remaining satellite programs. At this writing, in late 2018, the House Appropriations Committee has approved a budget for Earth science that is $116 million *higher* than NASA's request. This time, it did *not* specifically address the proposed cuts, thus opening a window for executive cancellation later on. If approved, these cuts will result in significant gaps in data about major climate-related processes, including data needed to verify emissions cuts programmed by the 2015 Paris Agreement.

Privatizing weather satellites and weather data

President Trump's NASA administrator is Jim Bridenstine, a former Republican Congressman from Oklahoma. The first elected official and non-scientist ever to hold that post, until very recently Bridenstine proclaimed denialist views. However, after his confirmation as NASA administrator Bridenstine told senators that his opinion had evolved. He now agrees with the scientific consensus that global warming is happening and that humans are its primary cause. Asked the reason for his changed views, Bridenstine said he "listened to a lot of testimony. I heard a lot of experts, and I read a lot" (Davenport 2018). He now plans to fully support NASA's climate science missions.

Bridenstine also supports the acquisition of weather and Earth science data from commercial satellites. The American Space Renaissance Act (ASRA), a bill he introduced in 2016, would direct the National Oceanic and Atmospheric Administration to revise its data provision rules to "ensure that [the agency] does not release more than the minimum amount of data required" under World Meteorological Organization Resolution 40. If passed, the ASRA would direct NASA to evaluate whether "industry can provide new or alternative architectures for Federal Earth science missions" and propose "new Federal programs that leverage such commercial capabilities, products, and services." This vision of hybridized government/commercial sourcing of Earth science data raises numerous questions. How will data quality standards be set and audited? How much data will be publicly released? What entity would preserve these data for the long term? Would scientists who need more than basic data now have to pay large, potentially prohibitive sums to acquire it?

Without careful work to ensure continuity and transparency, privatization could create significant discontinuities in weather data collection, preservation, analysis, and access, potentially reducing the quality and/or availability of climate data. As Gemma Cirac Claveras has observed, privatizing weather satellites challenges another long-standing scientific truce:

> a steady shift taking place in the perception of weather data, from part of a global commons to a global commodity to be bought and sold. What is at stake is the whole notion of 'the commons', the idea of recognizing and preserving a shared human and natural heritage through public institutions.
>
> *(Cirac-Claveras 2018)*

Peer review

As soon as he was appointed, the Trump administration's Environmental Protection Agency (EPA) director, Scott Pruitt, set to work deleting the phrase "climate change" from EPA web pages. By October 2017, this task had been largely accomplished. Yet this output-side attack received more attention than two other moves with potentially even more severe consequences for climate knowledge.

First, Pruitt decreed that scientists who receive EPA grants could no longer participate in EPA science advisory panels. Citing the Biblical prophet Joshua—"choose this day whom you're going to serve"—Pruitt declared that scientists could either receive EPA grants, or advise the agency on policy, but not both. (No similar injunction applied to industry-funded scientists, whom Pruitt has proceeded to appoint to EPA boards.)

Pruitt's policy reflects a longstanding theme of climate change deniers: that government research grants are a corrupting influence. The concept seems to be that scientists might seek to amplify the importance of problems they study, even to the point of conspiring to fabricate a global issue such as climate change, in order to keep bringing in grant money. Those who espouse this view rarely discuss the corrupting influence of private-sector money; indeed, many scientists cited by sceptics, and the think tanks that support them, have received substantial funding (including salary) from fossil fuel companies (Hoggan and Littlemore 2009). Pruitt's injunction appears paradoxical on its face, since those scientists most knowledgeable about any given regulatory issue are also most likely to seek research support from the EPA to study that issue. Pruitt's policy was immediately condemned by the American Association for the Advancement of Science (Leber 2017) and numerous other scientific societies.

Second, Pruitt repeatedly called for a "red team-blue team review" of climate science (Redfearn 2017). He characterized this exercise as an "honest, open, transparent discussion about this supposed threat to this country." The red team-blue team technique, taken from military and cyber-defense exercises, is designed to expose vulnerabilities that might be exploited by an attacker—not to calmly assess the quality of arguments and evidence using the best available expertise. A leaked email from Heartland Institute CEO Joseph Bast indicates that "folks at EPA" asked him for "recommendations" for "Red Team" scientists and economists, which Bast duly provided (Demelle 2017). Pruitt hoped to hold such an exercise on national television, a forum that strongly rewards personal charisma, hyperbole, and the reduction of complex issues to sound bites. As one climate scientist, a NASA veteran, put it:

> From my experience in both types of review, I can say confidently that red team-blue team exercises are not a mechanism for scientific debate. They are not designed to take a testable hypothesis and then look at whether observations and theory support or refute it. They are more like Heath Ledger's Joker in The Dark Knight, causing disruption, distortion and chaos.
>
> *(Rood 2017)*

Committed denialists of Pruitt's ilk ignore the decades of review and re-review of all major climate science conclusions. Indeed, the Intergovernmental Panel on Climate Change peer review process is probably the deepest and most thorough ever designed (Edwards and Schneider 2001; Farber 2007).

While the red team-blue team exercise remains on hold at this writing, the larger issue is Pruitt's redefinitions of conflict of interest and peer review. They target longstanding scientific truces, replacing them with a corrupt process under the guise of an audit culture. Should they take root throughout American science agencies, the nation's climate knowledge infrastructures will be severely damaged.

Conclusion

This chapter has examined specific challenges to routines and truces surrounding climate data. Starting in the 1990s, but accelerating dramatically after the turn of the millennium, routines and truces in the climate knowledge infrastructure became the targets of concerted attack. Mirroring trends in the larger society around it, an "audit culture" arose that sought to open the black boxes of scientific processes, seeking to expose illegitimately manipulated data and analysis. Under the banner of "transparency," it promoted the value of scrutiny from positions outside the established boundaries of disciplinary expertise, effectively cashing the chip of the "modest witness" whose virtue consists in having no direct stakes in the outcome. This position challenged prevailing norms of peer review, even to the point of disqualifying disciplinary experts as "government scientists."

The "audits" I discussed began as a check on scientific output, but moved on to scrutinize *inputs* to the knowledge-making process, including weather stations and raw data. All drew large contingents of vitriolic supporters hoping to witness a debunking of "AGW" (anthropogenic global warming). Yet instead of proving correct their suspicions of bias or illegitimate data adjustment, both surfacestations.org and BEST ended up *confirming* existing climate knowledge. BEST director Richard Muller changed his view on anthropogenic global warming as a result. The same cannot be said for the other two groups, which continue to contest data, methods, and the validity of peer review.

The most direct threat to climate knowledge infrastructures now comes from an American administration hostile to the very concept of anthropogenic climate change. It plans to eliminate key instruments, de-fund climate data analysis and other research, and rewrite the rules of peer review. It appears poised to privatize key satellites for collecting Earth science data, a move with complex, poorly understood ramifications for the future. These shifts target the chain of evidence regarding climate change, the stability of the peer review system, and the climate research funding paradigm in place since the 1950s.

The entire international climate infrastructure will suffer as a result. Nations with extensive scientific resources—especially the European Union, Japan, and China—may take up some of the slack. In 2018 the state of California announced plans to orbit its own carbon monitoring satellite, partially replacing a similar

NASA program deleted by the Trump administration. Scientists I know have discussed whether private-sector entities such as Google, Microsoft, Amazon, or Tesla—with access to privately-owned satellites, rockets, massive computing power, and "big data" methods—might join the climate knowledge enterprise, perhaps replacing lost government capacity. The only sure thing is that climate change will continue, whether or not our knowledge of it keeps pace.

Notes

1 Middle English "trewe" (sing.) and "trewes" (pl.), according to the OED.
2 Original caption: Smoothed reconstructions of large-scale (Northern Hemisphere mean or global mean) surface temperature variations from six different research teams are shown along with the instrumental record of global mean surface temperature. Each curve portrays a somewhat different history of temperature variations and is subject to a somewhat different set of uncertainties that generally increase going backward in time (as indicated by the grey shading). This set of reconstructions conveys a qualitatively consistent picture of temperature changes over the last 1,100 years and especially over the last 400 (Committee on Surface Temperature Reconstructions for the Last 2000 Years, Board on Atmospheric Sciences and Climate, and National Research Council 2006).
3 https://web.archive.org/web/20170713182820/https://wattsupwiththat.com/2017/06/14/wuwt-at-10-years-i-need-some-help-please/, accessed 15 December 2017.
4 National Centers for Environmental Information, www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/us-historical-climatology-network-ushcn, accessed 15 December 2017.
5 "Briggs on Berkeley's forthcoming BEST surface temperature record," WUWT blog entry, 6 March 2011, https://wattsupwiththat.com/2011/03/06/briggs-on-berkeleys-best-plus-my-thoughts-from-my-visit-there/, accessed 15 December 2017.
6 BEST land-surface temperature analysis, as shown to the US House Science, Space, and Technology Committee on 31 March, 2011, prior to peer review. Source: Richard A. Muller, "Statement to the Committee on Science, Space and Technology of the United States House of Representatives," available at science.house.gov/sites/republicans.science.house.gov/files/documents/hearings/Muller%20Testimony%20rev2.pdf.
7 http://climatechangereconsidered.org/, accessed 15 December 2017.

References

Arbesman, S. 2013. "Stop Hyping Big Data and Start Paying Attention to 'Long Data'." *Wired*. www.wired.com/2013/01/forget-big-data-think-long-data/

Argote, Linda. 1999. *Organizational Learning*. Boston: Kluwer Academic.

Aronova, Elena. 2017. "Geophysical Datascapes of the Cold War: Politics and Practices of the World Data Centers in the 1950s and 1960s." *Osiris* 32: 307–27.

Baker, Monya. 2015. "Over Half of Psychology Studies Fail Reproducibility Test." *Nature* 27. www.nature.com/news/over-half-of-psychology-studies-fail-reproducibility-test-1.18248

Bowen, Mark. 2008. *Censoring Science*. New York: Penguin.

Bowker, Geoffrey C. 2000. "Biodiversity Datadiversity." *Social Studies of Science* 30 (5): 643–83.

Bowker, Geoffrey C. 2005. *Memory Practices in the Sciences*. Cambridge: MIT Press.

Cirac-Claveras, Gemma. 2018. "The Weather Privateers: Meteorology and Commercial Satellite Data." *Information & Culture* 53 (3/4): 271–302.

Collins, Harry M. 1985. *Changing Order*. London: Sage.

Collins, Harry, and Trevor Pinch. 1993. *The Golem: What Everyone Should Know About Science*. Cambridge: Cambridge University Press.

Committee on Surface Temperature Reconstructions for the Last 2000 Years, Board on Atmospheric Sciences and Climate, and National Research Council. 2006. *Surface Temperature Reconstructions for the Last 2000 Years*. Washington, D.C.: National Academies Press.

Daniel, Howard. 1973. "One Hundred Years of International Co-Operation in Meteorology (1873–1973)." *WMO Bulletin* 22: 156–203.

Davenport, Christian. 2018. "NASA's New Administrator Says He's Talking to Companies About Taking Over Operations of the International Space Station." *The Washington Post*. www.washingtonpost.com/news/the-switch/wp/2018/06/05/nasas-new-administrator-says-hes-talking-to-companies-to-take-over-the-international-space-station/?noredirect=on&utm_term=.f4392946a0d5

Demelle, Brendan. 2017. "Heartland Institute 'Red Team' Climate Lists Revealed, And Science Deniers Are Upset with Pruitt." www.desmogblog.com/2017/10/25/heartland-institute-red-team-lists-revealed

Edwards, Paul N. 2010. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge: MIT Press.

Edwards, Paul N. 2013. "Predicting the Weather: A Knowledge Commons for Europe and the World." In *Cosmopolitan Commons: Sharing Resources and Risks across Borders*, edited by Nil Disco, and Eda Kranakis, 155–84. Cambridge: MIT Press.

Edwards, Paul N., Steven J. Jackson, Geoffrey C. Bowker, and Cory P. Knobel. 2007. *Understanding Infrastructure: Dynamics, Tensions, and Design*. Ann Arbor: Deep Blue.

Edwards, Paul N., and Stephen H. Schneider. 2001. "Self-Governance and Peer Review in Science-for-Policy: The Case of the IPCC Second Assessment Report." In *Changing the Atmosphere*, edited by Clark A. Miller, and Paul N. Edwards, 219–46. Cambridge: MIT Press.

Fall, Souleymane, Anthony Watts, John Nielsen-Gammon, Evan Jones, Dev Niyogi, John R. Christy, and Roger A. Pielke Sr. 2011. "Analysis of the Impacts of Station Exposure on the Us Historical Climatology Network Temperatures and Temperature Trends." *Journal of Geophysical Research* 116 D14120.

Farber, Daniel. 2007. "Modelling Climate Change and Its Impacts: Law, Policy, and Science." *Texas Law Review* 86: 1655.

Gitelman, Lisa. 2013. *Raw Data is an Oxymoron*. Cambridge: MIT Press.

Goldstein, Brett, Lauren Dyson, and Abhi Nemani. 2013. *Beyond Transparency: Open Data and the Future of Civic Innovation*. San Francisco: Code for America Press.

Hoggan, James, and Richard D. Littlemore. 2009. *Climate Cover-Up*. Vancouver: Greystone Books.

Houghton, John Theodore. 2001. *Climate Change 2001: The Scientific Basis*. Cambridge; New York: Cambridge University Press.

International Council of Scientific Unions. 2016. "ICSU World Data System." www.icsu-wds.org

Janis, Irving L. 1982. *Groupthink*. New York: Houghton Mifflin.

Jasanoff, Sheila. 1990. *The Fifth Branch*. Cambridge: Harvard University Press.

Jasanoff, Sheila. 2005. *Designs on Nature*. Princeton: Princeton University Press.

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Leber, Rebecca. 2017. "Scott Pruitt is Using the Bible as His Guide for Reorganizing EPA's Science Boards." *Mother Jones*. www.motherjones.com/environment/2017/10/scott-pruitt-is-using-the-bible-as-his-guide-for-reorganizing-epas-science-boards/

Lövbrand, Eva, Roger Pielke, and Silke Beck. 2010. "A Democracy Paradox in Studies of Science and Technology." *Science, Technology & Human Values* 36 (4): 474–96.

McIntyre, S, and R McKitrick. 2003. "Corrections to the Mann et al. (1998) Proxy Data Base and Northern Hemispheric Average Temperature Series." *Energy & Environment* 14 (6): 751–71.

McIntyre, S, and R McKitrick. 2005. "The M&M Critique of the MBH98 Northern Hemisphere Climate Index: Update and Implications." *Energy & Environment* 16 (1): 69–100.

Mann, Michael E. 2012. *The Hockey Stick and the Climate Wars: Dispatches from the Front Lines*. New York: Columbia University Press.

Mann, Michael E., Roger S. Bradley, and Malcom K. Hughes. 1998. "Global-Scale Temperature Patterns and Climate Forcing Over the Past Six Centuries." *Nature* 392 (6678): 779–87.

Mann, Michael E., Raymond S. Bradley, and Malcom K. Hughes. 1999. "Northern Hemisphere Temperatures During the Past Millennium: Inferences, Uncertainties, and Limitations." *Geophysical Research Letters* 29 (6): 759.

Mann, Michael E., Roger S. Bradley, and Malcom K. Hughes. 2004. "False claims by McIntyre and McKitrick regarding the Mann et al. (1998) reconstruction." www.realclimate.org/index.php/archives/2004/12/false-claims-by-mcintyre-and-mckitrick-regarding-the-mann-et-al-1998reconstruction/

Mann, Michael E., and Philip D. Jones. 2003. "Global Surface Temperatures Over the Past Two Millennia." *Geophysical Research Letters* 30 (15): 1820.

Menne, Matthew J., Claude N. Williams, and Michael A. Palecki. 2010. "On the Reliability of the US Surface Temperature Record." *Journal of Geophysical Research* 115 (D11).

Merton, Robert King. 1973. *The Sociology of Science*. Chicago: University of Chicago Press.

Miller, Clark A. 2007. "Democratization, International Knowledge Institutions, and Global Governance." *Governance* 20 (2): 325–57.

Muller, Richard A., Robert Rohde, Robert Jacobsen, Elizabeth Muller, and Charlotte Wickham. 2013. "A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011." *Geoinformatics & Geostatistics: An Overview* 1 (1). www.scitechnol.com/new-estimate-of-the-average-earth-surface-land-temperature-spanning-to-1eCc.php?article_id=450

Nelson, Richard R, and Sidney G Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge: Harvard University Press.

Office of Inspector General, National Aeronautics and Space Administration. 2008. "Investigative Summary Regarding Allegations That NASA Suppressed Climate Change Science and Denied Media Access to Dr. James E. Hansen, a NASA Scientist." Washington, D.C.: Office of Inspector General, National Aeronautics and Space Administration.

Oreskes, Naomi, and Erik M. Conway. 2010. *Merchants of Doubt*. New York: Bloomsbury Press.

Pearce, Fred. 2010. "Climate Wars: The story of the hacked emails." London: *The Guardian*. www.guardian.co.uk/environment/series/climate-wars-hacked-emails.

Redfearn, Graham. 2017. "EPA Chief Pruitt's 'Red Team' on Climate Science Is an Eight-Year-Old Talking Point Pushed by Heartland Institute." https://www.desmogblog.com/2017/06/13/epa-chief-scott-pruitt-red-team-climate-science-eight-year-old-talking-point-heartland-institute.

Rood, Richard B. 2017. "Red team-blue team? Debating climate science should not be a cage match." https://theconversation.com/red-team-blue-team-debating-climate-science-should-not-be-a-cage-match-80663

Roosevelt, Margot. 2011. "Critics' Review Unexpectedly Supports Scientific Consensus on Global Warming." *Los Angeles Times*. www.latimes.com/news/local/la-me-climate-berkeley-20110404,0,772697.story

Schatz, Gerald S. 1978. *The Global Weather Experiment: An Informal History*. Washington, D.C.: National Academy of Sciences.

Shapin, Steven, and Simon Schaffer. 1985. *Leviathan and the Air-Pump*. Princeton: Princeton University Press.

Smith, Marcia. 2017. "NASA'S FY2018 Budget Request Fact Sheet (updated September 24, 2017)." Arlington, VA: Space and Technology Policy Group. https://spacepolicyonline.com/wp-content/uploads/2017/03/NASA-FY2018-budget-request-Sep-24-2017.pdf

Spencer, Roy W., John R. Christy, and William D. Braswell. 2017. "UAH Version 6 Global Satellite Temperature Products: Methodology and Results." *Asia-Pacific Journal of Atmospheric Sciences* 53 (1): 121–30.

Stodden, Victoria, Jennifer Seiler, and Zhaokun Ma. 2018. "An Empirical Analysis of Journal Policy Effectiveness for Computational Reproducibility." *Proceedings of the National Academy of Sciences* 115 (11): 2584–89.

Watts, Anthony. 2009. "Is the US Surface Temperature Record Reliable?" Chicago: The Heartland Institute.

Weinberger, David. 2012. *Too Big to Know*. New York: Basic Books.

Wickham, Charlotte, Robert Rohde, Richard A. Muller, Jonathan Wurtele, Judith Curry, Don Groom, Robert Jacobsen, Saul Perlmutter, and Arthur Rosenfeld. 2013. "Influence of Urban Heating on the Global Temperature Land Average Using Rural Sites Identified from MODIS Classifications." *Geoinformatics & Geostatistics: An Overview* 1 (2). www.scitechnol.com/influence-urban-heating-global-temperature-land-average-using-rural-sites-identified-from-modis-classifications-vwBQ.php?article_id=588

3

AGAINST INFRASOMATIZATION

Towards a critical theory of algorithms

David M. Berry

We now live within a horizon of interpretability determined in large part by the capture of data and its articulation in and through algorithms. This novel space of experience and meaning creates a new envelope for economic valorization and leads to new forms of control and exploitation – and subsequently to new sites for social conflict. I want to argue that we can use critical theory for deepening our thinking about algorithms and data and understand how they manifest themselves in everyday life. Whilst it is not a perfect metaphor, it is clear that an understanding of a system of publishing requires more than studying just "books" in and of themselves (in as much as books might be understood as non-networked "websites" or data repositories). We also need to understand the systems and processes around the "bookish" objects and in particular their political economy. Similarly, with computational "data" objects, it is crucial that the algorithmic underpinnings are given critical attention – above and beyond what we might call its data logics, such as in "capta", databases, interfaces and streams – and connected to a wider political economy. Indeed, to think about data, and especially a "data politics" requires one to think across multiple levels of computational systems. To think about data, we have to think beyond data.

An algorithm is often defined as a series of rules followed to reach a computable end. The algorithm is a computational process followed, in a somewhat deterministic workflow, towards a conclusion. Algorithms tend to be defined through means-end, or instrumental rationality to perform specific functions and calculate specific results. Today we predominantly think of algorithms as the rules governing the operation of a computer program, but we need to remain alert to the conceptual utility of the concept of "algorithm". It seems to me that whilst we might deploy the concept of algorithm critically, it also has a tendency to obfuscate materiality due to its conceptual abstraction. So, we must be sure that the notion of the algorithm does not dis-embed computation, in effect reifying it.

I think this further raises the question of whether the term algorithm is itself ideological, in as much it conceals and distorts the legibility of computation (see Berry 2014). After all, the concept of "algorithm" often stands in for a variety of forms of computation: technical processes, software, code, source-code, etc. As an abstraction, it can obscure the specificity of computational instances, and may conflate all computational processes as "algorithmic". So, we need to unpack the concept of algorithm into specific material instances if we are to critically engage with computation (see for example Eubanks 2017, Noble 2018). But we must also be careful not to reduce algorithms to instrumentalism by implying that all agency rests only with human actors or social groups – algorithms are not merely tools.

It is the material specificity that is crucial to capture about computation above and beyond its sociological or cultural logics and hence its ability to structure a specific condition of possibility. This also applies when undertaking a critical analysis of data, which is similarly made up of a constellation of algorithms, protocols, systems and archives. Algorithms now shape and mediate our direct experience of a system of capitalist exploitation. In doing so algorithms utilize, in their material substructure, a number of what we might call fundamental material (or operational) concepts that frame software's operation. These material concepts are both operationalized and ideological. I want to argue that these concepts can be radicalized to perform a critique of the presently existing structure of society by an immanent critique. But we need to go further than focus on the issues raised by algorithms in and of themselves and think critically about the underlying system of domination that is immanent to computation and calculative reason in and of itself. We must interrogate and critique the will to power of algorithms and the political economy it makes possible (see, for example, Monahan 2018, Pasquale 2016, Srnicek 2017).

Today, in work, action and intellectual inquiry we see a growing use of computational systems to abstract, simplify and visualize complex "Big Data" phenomena and a tendency towards simplistic causal and statistical models to understand complex social and cultural phenomena, such as the notion of "social physics" (Pentland 2015). Chris Anderson has proclaimed the "End of Theory" as the data deluge has made the scientific method obsolete. Indeed, he claims that

> we can stop looking for models. We can analyze the data without hypotheses about what it might show. We can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find patterns where science cannot.
>
> *(2008)*

and that "with enough data, the numbers speak for themselves" (Anderson 2008). These claims reflect what we might call a "cult of data-ism" and a renunciation of the extended and important role of critical reason and theoretical thinking in modern society.

In contrast, I argue that a critical theory of algorithms is needed that prevents "algorithm studies" becoming merely an "academy of projectors". Rather, theory and its development are crucial to understand the contemporary situation through the confrontation of the object with its own concept. It is an approach that refuses to ignore and smooth over contradictions and contradictory claims and attempts to grasp the dynamic moment of the subject. By leaving open the possibility of a critical reflexive understanding of algorithmic history and tradition, it accepts the importance of the meaning structure of tradition but also seeks not idealize it. That is, that in computational societies, traditions might continue to embody interaction based on deception and distortion (in other words, ideology), and which can often be translated unreflexively into algorithmic forms. The cult of data-ism is a turn away from the project of seeking to understand society and culture through the application of critical reason in human affairs towards a data-deterministic world. It is problematic to erect an abstract and metaphysical standard by which human action and society can be judged – yet the cult of data-ism makes such a claim and works hard to produce and reproduce this new data-centric milieu. Whether history will be written by algorithms and data, even in its first draft through computational journalism, is subject to citizens' power to contest and challenge this new form of authority. One way to do this is by critiquing concrete examples of computationalism by drawing on a critical theory of algorithms. This enables us to challenge the cult of data-ism and an administrative approach to thinking about algorithms and instead to suggest different ways of being in a digital age.

In this chapter, I use these ideas to explore two sites for thinking about how one can begin to uncover a new politics appropriate to a contemporary computational milieu. The technical cannot be the locus of all politics, of course, as it could not cover all the varieties of political struggle expressed in society. But it is crucial that a politics of the computational is developed to examine the kinds of issues that are structured in and through data. To my mind there is little doubt that such a consideration requires a multidisciplinary approach to understand how algorithms are radically transforming politics, society, the economy and everyday contemporary life (Berry 2014; Berry and Fagerjord 2017).

I argue that a critical theory of algorithms (CTA) must be concerned with examining the particular historical conditions that give the present its shape in relation to the specific material and ideological formations that algorithms introduce into the social and economic conditions of society. A CTA seeks to explore human reality as a self-structured, self-unfolding and contradictory whole. Through this explication, reality becomes open to radical change, but nonetheless, a CTA does not claim to represent a complete picture of reality. In the context of computation this requires that we need to consider the specific historical ideas and practices within which we experience algorithms and in which they are made and remade. This means that we need to critique an ahistorical notion of the "algorithm" and critically interrogate metaphors and analogies that are necessary to explain but are not sufficient for understanding the instantiation of algorithmic forms in the new world of computation.

Here, I want to use the term algorithm, rather than software or code, not because algorithms capture more accurately the way in which the technical underpinnings of how digital technology are manifested through digital processes and systems, but because "algorithm" has become an increasingly important lens. We see it deployed as a concept in journalism and in academic work as a shorthand for a wide range of different computational systems. But equally, for this reason its use as a concept requires interrogation and historicizing. The digital world is not a static object; it is a highly dynamic and relational system that is in constant movement and undergoing continual change. For example, it is quite remarkable to note that the internet has never been taken off-line in order to be upgraded or changed, rather it is built through accretions and replacements that are slotted into or onto the existing system structure whilst it is still "running". This makes understanding the material specificity of algorithmic systems, like the internet, extremely important, and helps to show why an analysis that focused only on the "data" or "content" of the internet would be insufficient.

To deepen this analysis, I have previously introduced the notion of a laminated system for thinking about how one might undertake an analysis of computational systems, even the gigantic infrastructures of computation. Depending upon the level of analysis, different aspects of a computational system and its deployment of algorithms is made manifest. The layers I draw attention to include the physical, logical, codal, interactional, logistical and individuational layers (for a more detailed explanation of the notion of a laminated system see Berry 2014, 58). For example, we might look at the interface, the element in which reading largely takes place, through an analysis of the surface of computation a reading, which is deepened through a focus on what I call the "interactional" level. Algorithms are usually fixed into a specific layer – algorithms are implemented at specific levels to actuate particular functional requirements – and thus offer potential for comparative analysis between algorithmic levels. They can also transverse these layers and this helps to understand the systemic interaction that underwrites a computational system as a whole.

In this chapter I am particularly focused on the surface ("interactional") layer and how it functions in combination with the "logistics" layer responsible for organizing and distributing resources to create specific forms of work-processes and practices. The patterning of these layers of computation into vast laminated systems creates what I call *infrasomatization* (see Berry 2016). This notion draws on the work of Bernard Stiegler who has pointed to Alfred Lotka's and also Nicholas Georgescu-Roegen's notion of exosomatization as a crucial means of understanding computational capitalism (see Bobulescu 2015; Stiegler 2016, 95–96). Infrasomatizations are vast infrastructural configurations that create networks of cognitive agents to commodify human capacity for reason and thought – what Stiegler calls the *becoming-mnemotechnical* of every material, substance and product. I want to cast light upon these complex interactions through an immanent critique. By immanent critique, I refer to an approach drawn from the Frankfurt school, whereby the internal terms and concepts within a system are examined

in relation to the reality of the claims they make about and the actuality of the world. How computational systems are justified both discursively and in terms of their internal logics and the contradictions revealed in their organization. I want to interpret these algorithms by reflecting on how they structure subject positions within computational capitalism, and how humans are thereby affected by computation. I argue that this helps to demonstrate how algorithms are used as a technique to exercise power, and to produce strategic behaviour by shaping the labour, both physical and mental, of workers in specific digital environments. Indeed, the management of information and communication as an aspect of our everyday lives is one that is increasingly prevalent and the dynamics of these new hyper-individualized capillaries of power are seldom appreciated outside of academic and technical circles.[1]

We might note that in advanced capitalist societies, economic anarchy is interwoven with rationalization and technology to create fewer chances for mental and reflective labour. Under such conditions, the values of instrumental reason are accorded a privileged status since they are embodied in the concept of rationality itself. The confounding of calculation with rational thinking implies that whatever cannot be reduced to number is illusion or metaphysics. As a result, the conditions are created for a greater susceptibility of society to demagogic discourses and charismatic forms of power and a weakening of the potential for individuation. This forms part of the wider significance of a critical theory of algorithms and its contribution to social critique and critical thinking under contemporary conditions. Indeed, behind the ideological claims of data science and related approaches, particularly in Silicon Valley, this fetishism of calculation and computation is dominant. In spite of its efforts to reflect the object of analysis in terms of the manifest forms of development, such as here with algorithms, critical theory depends in its analysis on particular historical conditions. Therefore, it is crucial to maintain a dynamic distinction between social processes and resultant social forms by using commodity fetishism rather than the base/superstructure as its explanatory framework. Thus, institutional and ideological formations are not simple reflections of an economic base; instead, work has to be done to understand both culture and economy in relation to the growing use of computation.

The drive to use rationalisation and the insertion of algorithmic ways of doing and thinking permeates our everyday lives in contemporary computational societies. The introduction of measurable indicators of performance as standards of output and the monitoring and surveillance that computation makes possible are examples of such permeation. We also see the expansion of mediated experiences and systems of control in contrast to (dialogical) interaction in capitalist society. But equally seriously there is a lack of legitimacy for algorithmic systems that remain opaque and yet contribute to the structural problems associated with authority and legitimacy. As they continue to enmediate existing communications and media systems, algorithms' accelerating influence may generate systemic crisis and dangerous system failures. By enmediate I mean

> when a media form is no longer dominant, becoming marginal, and later absorbed/reconstructed in a new medium which *en-mediates* it. By using the term enmediate, I want to draw attention to the securing of the boundaries related to a format, that is, a representation or mimesis of a previous medium – but it is not the "same", nor is it "contained" in the new media. This distinction is important because at the moment of enmediation, computational categories and techniques transform the newly enmediated form.
>
> *(Berry 2013, 33)*

On this point, we might only reflect on the 2008–2013 financial crisis to see how rationalization and computation can create a heady mix in relation to profit-oriented corporations and individuals when the medium of communication is transformed. We might also note that increases in rationalization and computation can result in higher emotional intensity and promote irrational behavior (e.g. anger and outrage but also empathy and warmth). Similarly, the warnings from the 2016 US presidential election and the UK Brexit referendum show how computation and affect can be combined for political effect. This is a situation in which politics through the mediation of computationalism is rendered algorithmic. Critical theory responds by politicizing algorithms.

A critical theory of algorithms can offer an analytical framework for thinking about how society can escape the related fetishisation of newness and upgrade culture and contribute to a project of data politics (see Berry 2014; Berry and Dieter 2015). To do this I will look briefly at two case studies, one examining mental labour and the other physical labour. In the first, I look at the way in which social conflict is embedded within the machinery of algorithms and labour is transformed into a commodity through an interface. Here justificatory claims are linked to notions of "freedom" in a freelance "gig" economy mediated and performed by algorithms that are actually systems of control and exploitation. In the second, I look at how the interface-machinery dichotomy informs many attempts to discipline labour and enables radical attempts to objectify computation in the physical world. Here, the "digital market" of the Amazon Marketplace website is produced through a centralised computational system by making possible anti-human spaces through the use of "chaos algorithms" to proletarianise labour through a new form of radical objectification of the physical world through automation, rationalization and algorithmic systems. In this case a new algorithmic version of capitalist exploitation intensified by software and computation takes over from the successive forms of valorisation that dominated consciousness and society in the era of capitalist industrial production (Berry 2014).

"Freedom"

I want to suggest that one of the crucial approaches for thinking about algorithms and data is to attend to the way in which social conflict is submerged within the algorithmic form. A way to do this is to explore in more detail the new micro-labour

or micro-task practices, that is the new forms of labour that computational technologies make possible. This is where micro-task algorithms are claimed to be "literally mediating the future of work" (Bernstein et al 2015). The idea is to "draw on demand" labour as needed into a project from a "digital market place" – the highly visible examples being Uber, TaskRabbit and Upwork. This notion of micro-work is also connected to the notion of what has been called "flash teams" (Retelny et al 2014), which "leverage the scale of paid crowdsourcing for expert work". These systems allow one to "hire more people elastically in reaction to task needs, and pipeline intermediate output to accelerate completion times" (Retelny et al 2014).[2] In a sense this is the cooperation between brains that Stiegler (2010, 47) argues is "produced through grammatization systems which make possible the proletarianization of all those tasks conducted at the highest levels of nervous system activity". Crucially it is the real-time abstraction of labour-power as a potentiality, which we might think of as an unending stream of labour-power on demand in a similar fashion to an electricity or water supply that software and data make possible.

In the Fordist phase of capitalism, producing a stream of labour required the construction of massive factories to monitor and control labour through technologies such as the assembly-line. This was the era of industrial gigantism through the construction of huge, capital intensive factories with large work-forces gathered in a single location. Under conditions of computational capitalism, algorithms turn the factory inside out requiring the building of a distributed system of production reminiscent of the "putting-out" system as a means of subcontracting work. Also known as the "workshop" or "domestic" system, this method of production was considered a precursor to capitalism proper as industrialism. In this system workers used their homes or workshops to produce goods, allowing them flexibility in organizing their time and blending the notion of work and home. Similarly, under the micro-task systems, we see a return to this early capitalist form, but radicalized and intensified by the use of algorithms that simplify the difficulty of managing labour and production in such a decentralized system. These more recent experiments with micro-task production are nothing less than attempts to reinvent the world as a post-factory society. It requires the building of a new infrastructure of production by enframing labour-power within algorithmic "wrappers" that present a surface effect of a seemingly unending stream of abstract labour. This is akin to being within a videogame, with its own game mechanics, points awarded, badges, and stars, to create a sphere of production tailored to each individual worker who is encapsulated within the interface and thereby insulated from other workers and overseen and managed by the rule of algorithms.

This labour-power is made available via websites and apps creating a highly alienated form of labour-power that is disciplined and managed algorithmically through various forms of "signal" mechanisms, which are generated by the system, such as pay, ratings, reviews and metrics. In some sense, they are the realization of cybernetic systems that attempt to closely integrate machines and humans into tightly linked feedback loops and valorize labour through computation. This process requires management of bodies and minds and hence the creation of an

ideology of "freedom" through a creative economy but is really based on casualization, precarity and piece-work. The "boss" of the old factory is abolished by computation and replaced by the algorithm that guides, chides and informs through a personal device, such as a smartphone, whose very intimacy makes it compelling and trustworthy.

The paradigmatic example is Amazon Mechanical Turk (AMT), which Jeff Bezos, Amazon's CEO, calls "humans-as-a-service." As Bezos explains,

> normally, a human makes a request of a computer, and the computer does the computation of the task, but artificial intelligences like Mechanical Turk invert all that. The computer has a task that is easy for a human but extraordinarily hard for the computer. So instead of calling a computer service to perform the function, it calls a human.
>
> *(Bezos, quoted in Pontin 2007)*

AMT is named after the Mechanical Turk, an 18th-century chess-playing automaton actually driven by a small chess master hidden inside its case. The Mechanical Turk amazed contemporary observers with the seeming mechanization of the chess-playing skill. Meanwhile, in actuality the labour of playing chess was actually hidden behind the "interface" of the device. Thus, the skill of chess playing whilst contained within the system as a whole, was utilizing the embodied chess playing skills of a human. Similarly, today, computer scientists working on large-scale systems bracket off complexity by studiously ignoring how the functions they depend on are implemented – that is, they "black box" the components of the system design and worry about how they will be implemented at a later date. Programmers are taught to construct and respect "walls of abstraction" – functional modules that can be invoked in standard and consistent ways, hiding complexities within. The Amazon Mechanical Turk refers to itself as an "Artificial Artificial Intelligence", which aptly captures the way in which it conceptualizes its Human Intelligence Tasks (HITs) that "enable[s] technology builders to farm out massive volumes of small data processing tasks". The people requesting work ("requesters") "interact with the system primarily by posting tasks and receiving results produced by the pool of workers in the marketplace" who are paid a set piece-work fee (Irani 2015, 3). The tasks are usually small processing activities, such as labelling a photo, transcribing a fragment of text or performing a small calculation.

This notion of not only aggregating human beings through software, but also treating them as components or objects of a computational system is indicative of the kind of cybernetic thinking that is prevalent in computational society. In many ways this is a discretization of human activity, but it is also the dehumanization of humans through a computation layer used to mediate the use of social labour more generally. It also serves to show how the interface acts to reify social labour undertaken behind the surface, such that the machinery may be literally millions of humans "computing" the needs to the software, all without the user being aware of this. It also opens up the possibility of new forms of social labour, disconnected,

managed, controlled, monitored and disaggregated and re-aggregated on demand. In effect, people are selling "their idle brains to the companies and people who need the special processing power that they alone possess through marketplaces" (Pontin 2007). When operationalized into infrasomatizations by capitalist corporations, computation is in danger of creating a social shock, in as much as the new labour practices are often under the radar of labour laws and protections, but also hugely profitable and therefore have distorting effects upon the wider economy unless they are subject to regulatory control by governments and oversight by labour organisations, such as unions.[3]

Within computing the approach of using humans to stand in for algorithms is modeled as "Wizard of Oz prototyping". It allows a system to be broken down into constituent modules, which may involve processing of some kind but that might actually be fulfilled, usually temporarily by a human actor. However, if labour is cheap enough – or can be structured in such a way as to effectively massively reduce the cost, such as with the use of micro-piece work practices – then there may not need be a need to replace the human labour, which becomes a cog in larger computational systems. In effect, this allows corporations to tap virtually unlimited amounts of labour through aggressive task reduction to its smallest possible component – the creation of computational pin factories. These are the elements of an intensification of the division of labour made possible by algorithms, which are created and organized according to a logic of rationalization that does not address the social questions generated by this large-scale computerization. Unlike the manual labour usually associated with the division of labour, in this case it is the standardization, fragmentation and automation of cognitive mental work, usually understood as undertaken by white-collar workers who were previously largely immune from the labour management and proletarianization of blue-collar workers. Taken as a whole,this is a new infrastructure for mediating labour, interconnected across multiple, previously discrete computational systems, which link data, models and work into new aggregates.

Examining this "Wizard of Oz prototyping" approach, allows us to understand the process of constructing algorithms in a similar way to observing a process engineer working according to the principles of Taylorism, for example, using a watch to reify the relationship between worker and machine through standardisation and division of labour of work processes. These software products use "Wizard of Oz prototyping, a tried-and-true technique in HCI and AI . . . to put a human behind the curtain . . . until we understand how to engineer it". Indeed, these researchers claim that it is

> now possible to wire a wizard permanently into an interactive system . . . [which is] fully deployable from day one" (Bernstein et al 2010). Wizard of Oz prototyping involves a user "interacting with a computer system which is actually operated by a hidden developer – referred to as the "wizard." The wizard processes input from a user and simulates system output. During this process the user is led to believe that they are interacting directly with the system.
>
> *(Maudsley, Greenberg and Mander 1993)*

In reality, I argue that the interface acts as an ideological veil for the deployment of human labour. This form of prototyping is considered beneficial at early stages of the design cycle as it provides a means of studying and understanding user expectations and requirements. Maudsley et al argue that this approach is particularly suited to exploring design possibilities in systems that are demanding to implement, but here I am interested in how this practice becomes an end in itself as a means of utilizing what is conceptualized as abstract labour power, and which is indistinguishable to the user from an algorithm in and of itself. This is interesting to think about in relation to commodity fetishism as this is not the use of dead labour mediated through exchange, but rather the mediated use of living labour as a kind of algorithmic fetishism. Indeed, it appears as if the computer "magically" has the human intelligence, empathy, agency and skill to perform the tasks that are required – magnified if the laptop one is working on, or the phone in one's pocket is able to undertake skills that were previously thought of as only the domain of humans.

These approaches have fed into and made possible the new micro-labour practices that new computational technologies facilitate. As Bernstein explains,

> imagine a world in which the computing systems and applications that you use are not only far smarter than they are today but as smart as you and I are. And this is not a world that we have to wait 100 years for artificial intelligence to create but something that you could use tomorrow.
>
> *(Bernstein 2015)*

But of course, these systems are only partially computational with the "difficult" work undertaken by a hidden workforce of humans.

One example of this kind of design process is a plug-in for the Microsoft Word word-processor, christened Soylent, created by Bernstein et al and described as a "Word Processor with a Crowd Inside".[4] Soylent is a set of architectural and interaction patterns for "integrating crowdsourced human contributions directly into user interfaces" (Bernstein et al 2015, 85). In this case a prototype system for copyediting functions built on the Mechanical Turk system and obfuscated through a user interface incorporated into Microsoft Word.[5] It is a research project undertaken by Michael Bernstein and others in the MIT Human-Computer Interaction lab and therefore connects academic research and the growing phenomena of micro-labour practices. Here, workers' labour power is literally incorporated and mediated through the software. Soylent is part of a larger set of research projects around micro-task workflow management, but which also openly publish their results, including the source code of the system that can be compiled and run, allowing detailed analysis of the algorithm code, processes and interfaces. Soylent is, in the hyper-distributed nature of an "open access" society, also available open-source under the MIT license, and is hosted on Google Code.[6] However, these kinds of projects are also exemplary in showing the potential direction of these types of micro-task system and how easy they are becoming

to implement without any need to consider the normative questions raised by this form of anonymous crowdsourced labour. Indeed, these techniques are used to harvest human cognition processes for machine-learning data. For example, the use of image recognition training data through Google captcha and related interfaces (see Chew and Tygar 2004).

In the case of the Soylent software, the code is comprised of three main algorithms: (1) *Shortn*, a text shortening service that cuts selected text down to 85% of its original length on average without changing the meaning of the text or introducing new writing errors; (2) *Crowdproof*, a human-powered spelling and grammar checker that finds problems Microsoft Word is unable to identify, explains the error, and suggests corrections; and (3) *The Human Macro*, an interface for offloading arbitrary word processing tasks such as formatting citations or finding appropriate images. Each of these utilizes the Amazon Mechanical Turk system to distribute tasks to real-time human labour to undertake the text processing functions and check and anonymously return the results to the user (Bernstein et al 2010).

Let us take a closer look at this system as, even though it is a research project, it still allows us to understand and critique the ways of thinking engendered in the development of algorithmic systems. Perhaps the most noticeable is the assumed neutrality of the system but also the lack of a sociological understanding of the problem being investigated. This can often be manifested in a process of dehumanization of the subjects of the algorithmic system, but is also blind to race, gender, class and other social identities. This is done through a process of abstraction that strips particularity of the subject into data types, data streams and other structures that are then treated as disembodied "pure" labour-power.

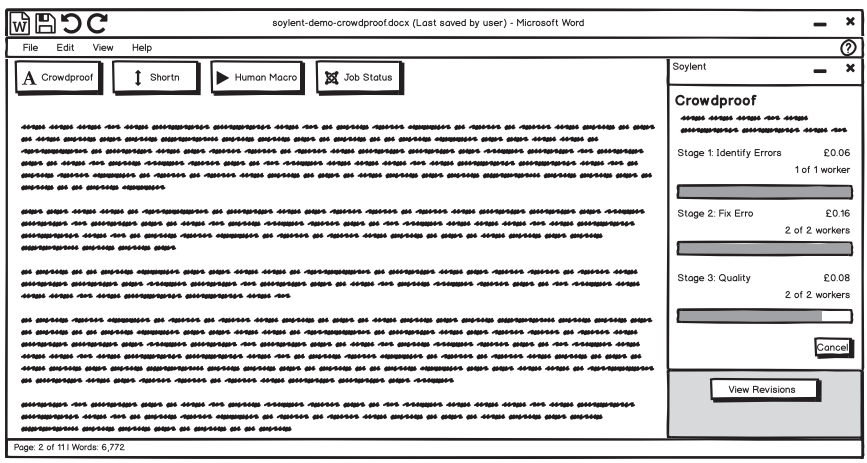


FIGURE 3.1 Soylent: a word processor with a crowd inside (Bernstein et al 2015)

So, for example, Soylent enables the user to automatically gather, employ, organize and discipline "crowds" of labour sourced via the Amazon Mechanical Turk system to undertake these word-processing functions through the human intelligence marketplace. This creation of an interface that abstracts away the underlying complexities of interacting with and using the Mechanical Turk workers changes the user experience of working with or interacting with human beings. As Irani notes in relation to the underlying Amazon Mechanical Turk (AMT) system, "AMT's interfaces render workers invisible, crowdsourcing entrepreneurs can imagine workers are in a better place. [There are a] proliferation of contradictory justifications for low wages on AMT – Turkers want fun, or live in 'developing' countries". Indeed, "with these stories, and by keeping low-status work at a distance, these professionals maintain the ideology of the non-hierarchical organization within their walls, keeping other kinds of new media work hidden behind the API or the interface" (Irani 2015, 16). The only relation users of the word-processor have with the workers whose labour they are using, is mediated algorithmically into a price-mechanism.

As mentioned earlier, even in the case of Soylent, which was nominally a research project, certain ideological constructions of workers and labour problems are apparent and have been implemented into the code, such as to discipline or control the workforce. For example, the designers developed what they call the "Find-Fix-Verify" paradigm in order to manage the labour and "quality" of the work of the "Turkers". As the researchers explain, the Find-Fix algorithm "forces Lazy Turkers to work on a problem of our choice. Allows us to merge work completed in parallel" and with the Verify algorithm "quality rises when we place Turkers in productive tension. Allows us to trade off lag time with quality" (Bernstein et al 2010, 58). The project introduces the Find-Fix-Verify design pattern as a general algorithm for programming/controlling crowds to complete open-ended tasks.

This is a crucial move and is why a critical theory of algorithms is urgently needed. Computation enables the creation of what may appear to be relatively benign algorithms (often abstractly called patterns), which not only can become property rights in and of themselves (as intellectual property rights) but also accentuate and encourage further development and abstraction of the exploitation of labour within algorithms. Indeed, the researchers make analogies with the popular UI technology Model-View-Controller, which has standardised programming user interfaces. In the case of Find-Fix-Verify, the algorithm functions to closely integrate labour to the needs of the user of the interface, in effect making the workers mere appendages of the machine. It is striking the way in which labour is inscribed into the system, but also de-humanized and reified into pure labour power, which is abstracted from its human form. But more than this, it is the purity of the algorithmic "pattern" that, stripped of normative content and decontextualized, appears to justify and encourage potentially exploitative and unjust labour practices. Here we need only note that these techniques are already used to

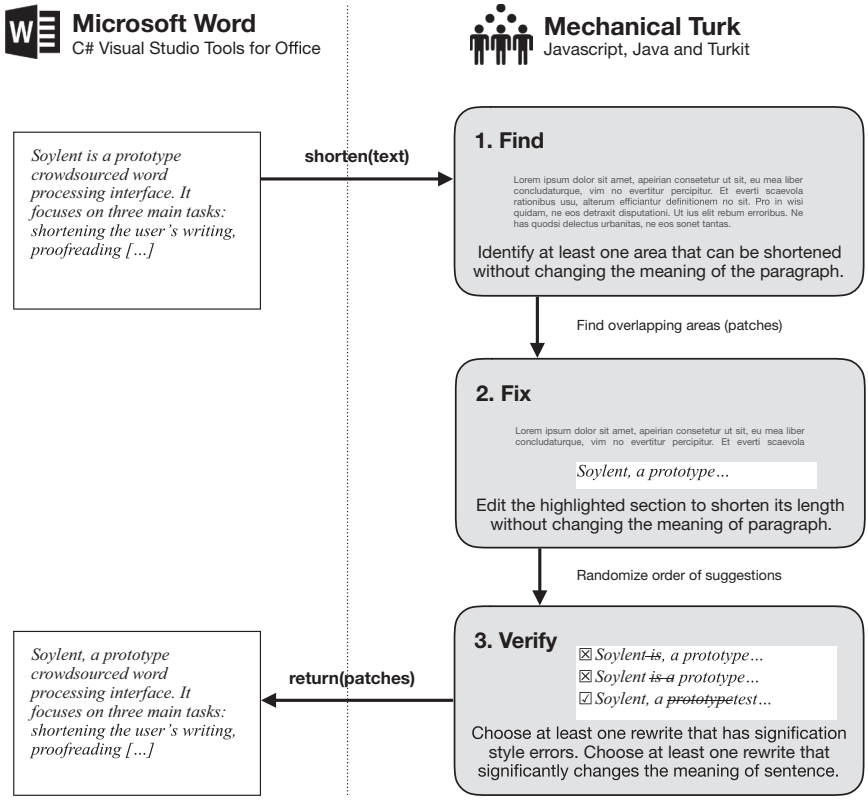


FIGURE 3.2 "Find-Fix-Verify" (Bernstein et al 2010, 58)

build platforms, such as ride-sharing apps (e.g. Uber, Lyft), delivery services (a.g. Deliveroo, UberEATs), domestic workers (e.g. Taskrabbit, Postmates) and many other new platforms.

This new paradigm to "stream" labour-power, points towards the possibility of a new class of crowd-powered interfaces through a proof-of-concept prototype. Here the machinery is obfuscated by the interface that may be human labour, an artificial intelligence (AI), or a standard algorithm. From the perspective of the user of the interface there is no difference. This procedure reifies the relationship and creates a Command-Execute relationship between user and underlying process. It foregrounds the ends of the process and obfuscates the means, regardless of the exploitative relationship it may be encoding or making possible. It goes without saying that the social and political consequences of the rationalization and abstraction of real human beings within computational systems is thereby minimized and social conflict is submerged into software.

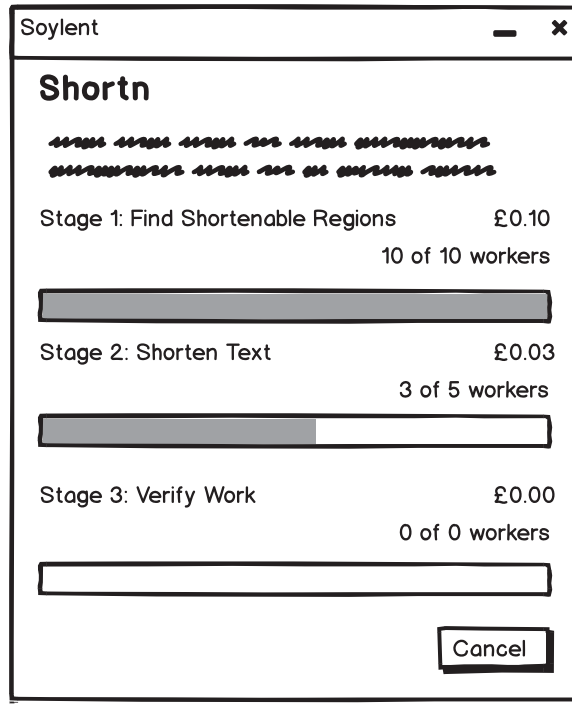


FIGURE 3.3 Shortn algorithm: even though it claims use of a "wizard" it nonetheless informs the user of the current cost and number of workers currently working

Perhaps this software class should therefore be better thought of as reifying labour into algorithmic units – actually called "blocks" in Retelny et al (2014) – but what we might describe as "human-labour blocks". This type of software pattern allows the user to become a "manager" of teams of people in a project even if they may not actually be aware that they are managing people due to the distantiation created by the mediation of the algorithms. Retelny et al (2014) describes the software package Foundry, for example, as a flash team authoring environment with "strongly typed handoffs" and "support for diverse expertise, and a runtime management platform".[7] That is, the software simplifies the process of managing diverse teams that are constituted from micro-task systems. These systems, rather than try to recreate the strengths of "in-person expert teams", create the conditions for a "beyond being-there" vision of expert crowd work, that enables remote-working micro-task type short-term precarious work. In other words, hiring and firing workers is now possible seamlessly and invisibly within the space of an interface.

This is an example of where computation serves to hide social labour such that workers are hidden "behind web forms and APIs [which] helps employers see themselves as builders of innovative technologies, rather than employers

unconcerned with working conditions" (Irani and Silberman 2013). As Marcus Courtney of WashTech, a technology workers union says, "what Amazon is trying to do is create the virtual day laborer hiring hall on the global scale to bid down wage rates to the advantage of the employer" (Mieszkowski 2006).

Attempts have been made to contest these systems and the algorithmic logics they instantiate. For example, as a response to the difficulty of collective action, unionisation or political mobilisation against various forms of human intelligence technology systems, which are built to deliberately obfuscate and limit the formation of a social group or identity. Salehi et al (2015) built "Dynamo". Dynamo was designed as a platform to support the Mechanical Turk community in forming publics around issues and then mobilizing them (see Harris 2014). Whilst laudable, it is interesting to see how the problem of the reification of human labour through mediating technologies is posed as a problem to be solved through a new set of technologies to remediate these relationships – technology is offered as a response to the problems of technology. However, these new forms of organizing are still nascent, and the Dynamo system has been largely dormant since its inception – perhaps pointing to the limitations of this kind of somewhat anti-political solution to worker organization and community building.

Some Turkers have formed collectives such as TurkerNation, MTurkGrind, and Reddit's/r/HITsWorthTurkingFor (Human Intelligence Tasks) using social media, etc. Amazon Mechanical Turk's legal terms create strong and often unfair power relationships between works and employers ("requesters") "enabling requesters to pay for data and nothing more. Requesters have full discretion to deny payment without justification to workers or to Amazon . . . requesters do not pay to train and maintain employees and infrastructure" (Irani 2015). Indeed, each worker is rated using metrics based on "how many of their tasks have been approved or rejected. The most productive are invited to become 'Masters' and gain exclusive access to better-paying tasks" but Amazon can also "deactivate Turkers' accounts at any time, without giving a reason." The workers themselves are often cognizant of their new subject position inside computation systems, for example, one Turker wrote on the Dynamo site, "I am a human being, not an algorithm, and [employers] seem to think I am there just to serve their bidding." (Harris 2014). The future of these platforms and the labour relations they create is symptomatic of the wider proletarianisation and automation that society faces under this technical imaginary. However, conversely we might also see the political potential for a powerful inversion of the usual idea of computation. Algorithms are increasingly associated with agency, value and freedom whereas worker's agency is to be automated away. By counter-intuitively self-identifying as an algorithm, a worker might assign this new agency and value to herself, within a computational milieu but not completely controlled by it. By negating a particularly capitalist notion of labour made possible in the new data capture and the algorithmic processing capacities of streams, new forms of political subjectivity are made possible. If the worker were to understand her own labour as algorithmic before the enframing logic of capitalism alienates it, that algorithm, that labour, is seen as a product of the worker rather than the system itself (see Berry 2011, 142–171).

"Chaos"

Next, let us look at the use of algorithms within Amazon's most advanced warehouses. In Amazon's physical stores, multiple objects are packed tightly into a warehouse space, which is computationally managed through a complex technology stack. What is interesting here is that, in contrast to the previous example, Amazon sees a commercial advantage in removing human labour and human cognition out of the system and replacing it with full-scale automation as much as practically possible. Of course, in this extreme case of warehouse automation it is its systemic implementation that is of interest. In other words, external to the algorithmic warehouse there still needs to be the technical a priori of human technicians, programmers, engineers and so forth to keep the warehouse running, but internally the system is a completely automated process that is organized computationally.[8] This use of computation to "bracket" space and create autonomous zones is highly suggestive of the manner of computational implementation more generally, and of the need to demarcate boundary zones, gatekeepers and boundary objects that enforce this dichotomy between what we might call real-space and algorithmic-space.[9]

Here computation is used to rethink the question of space, volume and organization in an extremely radical way in order to reduce the reliance on physical labour-power. Thus, the optimization of space internal to this system is done computationally: Amazon knows the exact dimensions of every product in its warehouses and the exact dimensions of vacant shelf space as well as the entire warehouse capacity and uses a set of algorithms to match object-space to the algorithmic-space of the warehouse. Amazon has even created a gestural model of the human body for situations where labour is still unavoidable for so-called human "pickers". Objects are retrieved using computer-controlled robots (formerly Kiva-branded orange robots) gliding swiftly and quietly around the warehouse bringing forth the stacks to the front of the building. These robots then glide the storage stacks back to the most efficient places, depending on frequency of access and contents of the stacks, rather like the way in which computers hold frequently accessed memory contents physically closer inside the processor. From the outside to humans, this system looks extremely disorganised and illogical. In fact, it represents the objectification of what Amazon calls a *chaotic storage algorithm*. The warehouse is in effect a reification of the code into the materials of stone, metal, plastic and human labour. The result is that the system functions at the highest rates of efficiency in the retail industry, with (some) humans removed from dealing with thinking about specific aspects of storing, retrieving (stowing/picking) objects to the greatest possible extent. Indeed, algorithms are deployed to manage the movement, organisation and cognitive-processing of the remaining workers too. As the system works, full-scale data collection enables new models to be created, workers to be tracked and new optimizations to be discovered and implemented.

It is striking that Amazon uses this so-called "chaotic storage" algorithm that optimizes storage through mediating databases and which writes onto the physicality of the warehouse building, stacks and remaining human labour by prescribing the

organization, movement, practices and gestures. For example, if Amazon receives a shipment of 500 copies of a specific book, they do not necessarily store the 500 copies in one location together. Instead, they can distribute the books to different areas of empty shelf space and their location is recorded in the database because the capacity for thinking about organising and storage is delegated to algorithms. The mediation of archival strategies, storage, and its materiality through the operation of robots and specialised storage units enables further optimizations, complexities and computational logics.

It is this process of objectification that once again demonstrates the need for a critical theory of algorithms. The recasting of the material world into the shapes dictated by computational analysis or computational processes (e.g. transduction) creates an absolute alienation of labour.[10] Using these augmented algorithms a human picker who has been working for years in the warehouse is no more productive than one who has worked for a week, since her familiarity with the layout of the warehouse is irrelevant. The only thing that matters is how quickly she can move her hands and eyes from point A to B, as directed by the algorithm using the data scanned in when products reach the warehouse. The amount of training required by new employees is therefore remarkably lower when using chaotic storage algorithms. It is not necessary for workers to memorize the entire warehouse layout or even single storage locations. This allows Amazon to potentially replace staff more easily or hire seasonal workers during peak times as humans do not even have to think about where to move their hands, as an algorithm-controlled laser points to the correct shelf and the order of picking items. All of this recasts the material world and labour-power into the shapes dictated by computational analysis or computational processes. This also highlights that algorithms are deployed not just to aid organizational efficiency, but also to deskill, proletarianize and pacify the human components of these algorithmic systems. Further integrating workers into the streams of algorithms as partial objects rather than sensual human beings, through new infrasomatizations.

However, it is precisely here in the notion of chaos as a model for the organization of labour and production that another world is visible. The suggestive idea of a "chaotic" system that operates at a higher level of efficiency than a traditional bureaucratic hierarchical warehouse might be deployed to negate the existing standards of dehumanized workplace organization. Of course, the Amazon model of implementation of deskilling is not to be applauded in and of itself, rather the idea of an "organized anarchy" has suggestive political implications. It prompts the asking of questions around the subordination of the worker to the job, the creation of mere appendages to the machine and the fetish of the assembly-line as a model of disciplining workers. Instead, an organized anarchy of algorithms, as democratic self-organization, might link the resources available and their distribution to a new system of needs.

Conclusions

I have sought to show how the study of actually existing algorithmic systems can contribute to developing and deepening a critical theory of algorithms.

By linking the strategies of thinking about questions of data power and politics to the project of a critical theory of algorithms a more sophisticated approach to offering a critique of computational capital is made possible. By situating its analysis at the mediation point between algorithm and labour, the relations of power and reification through data, the algorithmic form as a site of power is brought forward more starkly. One of the idiosyncrasies of computational or algorithmic capitalism is its widespread adoption of software development processes like "open source" (see Berry 2008). This means that there is a tendency to use widely available non-proprietary technologies in innovative combinations to create new technology stacks. But this also allows that these technologies can be "read" as their underlying logics are often available as source code offering a potential for immanent critique. This presents critical approaches with a tremendous opportunity to study, explore and understand these systems and their algorithmic structures to enable theorizing and to challenge their normative and often exploitative forms. It also points to the importance of seeking to read proprietary algorithms in situations where the social implication of softwarization have pernicious outcomes.

With the rise of Big Data, data itself becomes another important resource for input into these systems. Clive Humby has described a kind of process where

> data is the new oil . . . Data is just like crude. It's valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc. to create a valuable entity that drives profitable activity; so must data be broken down, analyzed for it to have value.
>
> *(Palmer 2006)*

Or as *Wired* put it,

> like oil, for those who see data's fundamental value and learn to extract and use it there will be huge rewards. We're in a digital economy where data is more valuable than ever. It's the key to the smooth functionality of everything from the government to local companies. Without it, progress would halt.
>
> *(Toonders 2014)*

This extractive metaphor, which is rich in illustrative description, is limited for understanding the processes of creating, maintaining and using data, and thus needs to be unpacked and critiqued. Indeed, seeking to open the databanks seems to me to still be an important tactic in a data politics, also the opacity of algorithms need to be questioned and linked to a wider political struggle.

These new systems create imbalances of power. Amazon does not "set minimum rates for work, which can pay less than $2 an hour, and takes a 10% commission from every transaction. Employers can even refuse to pay for work altogether, with no legal consequences" (Harris 2014). The way in which the organization of society is crystallized in algorithms and the institutions and networks they make

possible need urgent and better methods for understanding and exploring their constitution, capacities and implications. Equally, we need to undertake more work to understand the mechanisms that determine the infrastructures of algorithms, which themselves may be computationally mediated and have an ideological effect. The processes of algorithm production, reproduction, distribution, exchange and consumption are complex and multi-layered, as a result a set of critical methods for understanding these new tools, mechanisms and procedures need to be further refined to capture this complexity. This chapter has sought to be a contribution to the project of data politics by contesting the invisibility of algorithmic infrastructures, critiquing infrasomatizations as new cultural forms whether as algorithms, apps, Big Data or machine-learning. As algorithms disappear from view into these vast new infrastructures, the need for social critique becomes pressing. Identifying the pacifying effects of algorithms and their exploitation of labour-power offers a critical lever to crack open the black-box of computational capitalism.

Notes

1 These issues are explored in depth in the work of Irani (2015), who focuses on how social conflict is mediated through particular assemblages of algorithmic systems.
2 We should be clear that this form of crowd-sourcing of human labour-power, often low-paid or even unpaid, is not limited to the Web 2.0 economy of Silicon Valley. Indeed, academic research projects have in some instances tried to utilize the "wisdom of crowds" to make up for the low amounts of funding available in academia, a perceived "democratic" or "participatory" advantage of these approaches for impact in the humanities, or a lack of local expertise. In whatever form these systems are deployed, there remain significant questions to be addressed in relation to issues of exploitation, reification and a duty of care towards one's crowd-sourced "workers."
3 It is likely that we will see future social conflicts generated through greater use of automation systems that function in this way. In effect they create a sense of algorithmic fetishism in that provided the labour is hidden behind algorithms, then the programmer/user who interacts via dashboard interfaces and computer programming code need not acknowledge its social character. For the worker on the other side of the interface, the demands on agency, physical labour and emotional control and self-disciplining are likely to create severe psychic tensions in terms of pendulum swings between anomie and fatalism. A single bad review or rating from a customer or client can instantly cause termination of the employment – the reasons for which are seldom given. Whether this will create the conditions for a politics of necessity and a sociological reflexivity remains to be seen, but will be a potent future source of labour disenchantment beyond that of the traditional working class.
4 See http://projects.csail.mit.edu/soylent/
5 Other examples include the software Foundry an "end-user authoring platform and runtime manager. Foundry allows users to author modular tasks, then manages teams through handoffs of inter-mediate work." (Retelny et al 2014). This allows the user to coordinate isolated contractors, and micro-task crowdsourcing techniques from online marketplaces such as Upwork. Also, Dynamo a platform to support the Mechanical Turk community in forming publics around issues and then mobilizing. It structures "labor to maintain efforts forward motion" and prevent the "twin perils of stalling and friction" (Salehi 2015).
6 https://code.google.com/p/soylent/
7 See https://hci.stanford.edu/publications/paper.php?id=284

8 These algorithmic-spaces also require the condition of possibility created by environmental technologies, buildings, corporate structures, and so on.
9 Elsewhere I have used the distinction of compute-computing and compute-computed to think through this method of producing algorithmic zones through machine-learning (see Berry 2017).
10 This is not just taking place in the market economy, for example, the logic of the British Library's National Newspaper Building, Boston Spa, United Kingdom, which opened January 23, 2015. This is a facility that stores 33 linear kilometres of newspapers, 290,000 bound volumes. Here, the temperature is a constant 14 degrees centigrade, and 55% humidity to lengthen the life of fragile newspaper, oxygen levels in the void will be reduced to 14.6% to eliminate the risk of fire (air is 20.95% oxygen). In the BL Newspaper store newspapers are stored in high-density racking 20 metres high and collection items are retrieved by robotic cranes, which transfer stacks of newspapers via an airlock to a retrieval area. Indeed, other examples include: North Carolina State University James B. Hunt Jr. Library automated book retrieval system ("BookBot"), University of Missouri, Miller Nichols Library automated book retrieval system ("RooBot"), the five-story underground robo-library at the Joe and Rika Mansueto Library at the University of Chicago, and Macquarie University Library (Australia) – automated storage and retrieval system.

Bibliography

Anderson, C. 2008. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete." *Wired*, accessed 18/12/2015, http://www.wired.com/science/discoveries/magazine/16-07/pb_theory.

Ars Industrialis. n.d. Organologie, accessed 20/03/2016, http://arsindustrialis.org/vocabulaire-organologie.

Bernstein, M., Little, G., Miller, R. C., Hartmann, B., Ackerman, M. S., Karger, D. R., Crowell, D. and Panovic, K. 2010. "Soylent: A Word Processor with a Crowd Inside." In *Proc. UIST 2010*. New York: ACM Press.

Bernstein, M. S., Little, G., Miller, R. C., Hartmann, B., Ackerman, M. S., Karger, D. R., Crowell, D. and Panovic, K. 2015. "Soylent: A Word Processor with a Crowd Inside." Communications of the ACM, August 2015, 58(8), 85–94.

Berry, D. M. 2008. *Copy, Rip, Burn: The Politics of Copyleft and Open Source*. London: Pluto Press.

Berry, D. M. 2011. *The Philosophy of Software*. London: Palgrave.

Berry, D. M. (2013) Against Remediation, in Lovink, G. and Rasch, M. (Eds.) *Unlike Us Reader: Social Media Monopolies and Their Alternatives*. Amsterdam: Institute of Network Cultures.

Berry, D. M. 2014. *Critical Theory and the Digital*. New York: Bloomsbury.

Berry, D. M. 2016. "Infrasomatization." *Stunlaw*, accessed 1 Dec 2018, http://stunlaw.blogspot.co.uk/2016/12/infrasomatization.html.

Berry, D. M. 2017. "Prolegomenon to a Media Theory of Machine Learning." *Media Theory* 1(1). http://sro.sussex.ac.uk/70336/.

Berry, D. M. and Dieter, M. 2015. *Postdigital Aesthetics: Art, Computation and Design*. London: Palgrave Macmillan.

Berry, D. M. and Fagerjord, A. 2017. *Digital Humanities: Knowledge and Critique in a Digital Age*. Cambridge: Polity Press.

Bobulescu, R. 2015. "From Lotka's Biophysics to Georgescu-Roegen's Bioeconomics." *Ecological Economics* 120, 194–202.

Chew, M. and Tygar, J. D. 2004. "Image Recognition CAPTCHAs." *Proceedings of the 7th International Information Security Conference (ISC 2004)*, Springer,

pp. 268–279, accessed 01/12/2018, https://people.eecs.berkeley.edu/~tygar/papers/Image_Recognition_CAPTCHAs/imagecaptcha.pdf.

Eubanks, V. 2017. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. London: St Martin's Press.

Harris, M. 2014. "Amazon's Mechanical Turk Workers Protest: 'I am a Human Being, not an Algorithm'." *The Guardian*, accessed 04/02/2016, www.theguardian.com/technology/2014/dec/03/amazon-mechanical-turk-workers-protest-jeff-bezos.

Irani, L. 2015. "The Cultural Work of Microwork." *New Media & Society* 17(5): 720–739.

Irani, L. C. and Silberman, M. S. 2013. "Turkopticon: interrupting worker invisibility in Amazon mechanical Turk." (CHI 2013 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, April 27–May 2, 2013), New York, ACM, 2013: 611–620, accessed 07/10/13, http://wtf.tw/text/turkopticon.pdf.

Maudsley, D., Greenberg, S. and Mander, R. 1993. "Prototyping an Intelligent Agent through Wizard of Oz." *Interchi 93 Conference Proceedings*, 277–284.

Mieszkowski, K. 2006. "I make $1.45 a week and I love it." *Salon*, accessed 04/02/2016, www.salon.com/2006/07/24/turks_3/.

Monahan, T. 2018. "Editorial: Algorithmic Fetishism." *Surveillance & Society*, 16(1): 1–5.

Noble, S. U. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.

Palmer, M. 2006. "Data is the New Oil." *ANA Marketing Mastros*, accessed 01/12/2018, http://ana.blogs.com/maestros/2006/11/data_is_the_new.html.

Pasquale, F. (2016) *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press.

Pentland, A. 2015. *Social Physics: How Social Networks Can Make Us Smarter*. London: Penguin.

Pontin, J. 2007. "Artificial Intelligence, With Help from the Humans." *The New York Times*, accessed 15/01/2016, www.nytimes.com/2007/03/25/business/yourmoney/25Stream.html?_r=0.

Retelny, D., Robaszkiewicz, S., To, A., Lasecki, W. Patel, J., Rahmati, N., Doshi, T., Valentine, M. and Bernstein, M. S. 2014. "Expert Crowdsourcing with Flash Teams, In UIST." ACM Symposium on User Interface Software and Technology, accessed 13/12/2015, https://hci.stanford.edu/publications/paper.php?id=284.

Salehi, N., Irani, L. C., Bernstein, M. S., Alkhatib, A., Ogbe, E., Milland, K. et al. (2015). "We Are Dynamo: Overcoming Stalling and Friction in Collective Action for Crowd Workers." *UC San Diego*. http://dx.doi.org/10.1145/2702123.2702508.

Srnicek, N. 2017. *Platform Capitalism*. Cambridge: Polity.

Stiegler, B. 2010. *For a New Critique of Political Economy*. London: Polity.

Stiegler, B. 2016. "The New Conflict of the Faculties and Functions: Quasi-Causality and Serendipity in the Anthropocene." *Qui Parle: Critical Humanities and Social Sciences*, 26(1): 79–99.

Taffel, S. 2015. "Perspectives on the PostDigital: Beyond rhetorics of progress and novelty." *Convergence: The International Journal of Research into New Media Technologies*, accessed 11/12/2015, http://con.sagepub.com/content/early/2015/01/22/1354856514567827.abstract.

Toonders, Y. 2014. "Data is the New Oil of the Digital Economy". *Wired*, accessed 01/12/2018, www.wired.com/insights/2014/07/data-new-oil-digital-economy/.

4

SURVEILLANCE CAPITALISM, SURVEILLANCE CULTURE AND DATA POLITICS[1]

David Lyon

Introduction

A surveillance scandal involving Facebook exploded in 2018. In 2015 a political consulting company, Cambridge Analytica (CA), specializing in influencing voters, obtained access to personal data mined from 87 million Facebook users (Davies 2015). A Cambridge University social psychologist named Aleksandr Kogan built an app to harvest data from unwitting Facebook users. They were asked to take a survey from which psychological profiles were constructed and intended to predict their behaviour. The users were unaware that the data would gain access to their friends, or that another company, CA, was involved.

It was revealed that 270,000 Americans took the survey, enabling Kogan and his colleague, Alexander Nix, to develop a model predicting the personalities of all adult US citizens, that was then passed to CA. It is unclear which data were used, but CA worked for Ted Cruz and then for Donald Trump in the 2016 presidential election campaign. Steve Bannon, who was to be Trump's White House Chief Strategist for the first seven months of his presidency, was on CA's board and Robert and Rebekah Mercer, Republican Party supporters, backed CA financially. As for Facebook users, they were unaware that data from them and their friends were being used for these purposes.

The news about these activities did not fully break until 2018, when a Canadian, Chris Wylie, who used to direct research at CA, turned whistleblower and informed *The Guardian*, which on March 21 published details of what had happened, based on documents from CA. This prompted government hearings in the UK, USA – featuring Facebook's founder, Mark Zuckerberg among others – and in Canada in April 2018. And it also generated a storm of outrage and interest in Facebook's activities, particularly over privacy of users' data, that had repeatedly been the subject of controversy almost since Facebook was founded. A #deletefacebook hashtag

appeared, attracting much attention, and inquiries about how to remove oneself from Facebook grew rapidly, especially in the UK and Canada.

All this represents a new departure for studies of surveillance, an important marker of something that has been simmering under the surface for a number of years but finds concrete expression in the 2018 Facebook scandal. The key to this is that the internet is a surveillance space that is inherently fluid, liquid (Bauman and Lyon 2013). Such liquidity tends to blur boundaries, flowing across previously assumed activities and categories. For some time, for instance, the categories "online and offline" have seemed less and less salient to how people actually spend their daily lives. While these refer to distinct experiences – touch and smell, for example, are not yet available online – much of life is in fact lived "on the internet," in almost constant contact with, or finding out about, others who are not physically present. The latter category is especially interesting, because not only do people encounter and experience online surveillance, they also engage with it. This is "social surveillance" (Marwick 2012).

This happens because the internet has become a surveillant space that also smudges the distinctions between monitoring and tracking activities of security agencies, police and corporate marketers and advertisers on the one hand, and the surveillance initiatives of everyday life, on the other. What security agencies, police and corporate marketers do is hard to discern, for a number of reasons, including agency and commercial secrecy. But everyday surveillance is not well researched yet, either. Finding out about others, or "social surveillance," has many faces, from the relatively benign searches for classmates or potential romantic partners, to surveillance of groups and individuals that some wish to "name and shame" through forms of "digital vigilantism" (Trottier 2017).

Surveillance data are thus key to the functioning of the internet; they are part of what constitutes the internet at every level. They make possible many activities, both those that become visible in public scandals such as that affecting Facebook as well as those that are as yet relatively unknown. The internet, including surveillance data, also facilitates debates over surveillance activities, and over data themselves, thus also becoming an intrinsic dimension of the politics of surveillance, and of the internet itself. It is some of these complex inter-relations between the surveillance that characterizes large global organizations and surveillance involving the mundane activities of everyday life that now have to be explored if contemporary kinds of surveillance are to be understood. An important question is this: under what circumstances are the politics of data normalized or radicalized? And how do ordinary users' practices make a difference?

To paint with a broad brush, I shall frame this discussion in terms of two wide-ranging concepts, surveillance capitalism and surveillance culture. The first is associated with scholars such as Shoshana Zuboff and Mark Andrejevic and the latter with figures such as Alice Marwick and Anders Albrechtslund. I have also contributed to this research enterprise (Albrechtslund 2008, Lyon 2017, 2018, Marwick 2012, Zuboff 2015). Both surveillance capitalism and surveillance culture depend on data but often in different ways and with different consequences. I shall show

that surveillance capitalism is the source of the systems that enable many aspects of surveillance culture, and that at present much that counts as surveillance culture is supportive of surveillance capitalism. But this is not inevitable, as evident in the case of the Facebook scandal of 2018. The conditions of possibility – surveillance data in this case – do not produce predetermined outcomes. Or so I shall argue.

Surveillance capitalism

Let me turn to the first topic of the duo, surveillance capitalism. To focus on surveillance capitalism is to note the ways that surveillance is moving more rapidly towards centre stage in the political economy of the early twenty-first century. It is to grasp the immense power and profitability of personal data and to see why not only corporations but government departments, health-care systems, educational establishments and of course policing and security initiatives are so eager to follow the Big Data bandwagon into new realms of user-transparency, efficiency, productivity and power.

Facebook is a prime example of a surveillance capitalism corporation. What came to be called social media was in its infancy at the century's turn. Friendster, founded in Kuala Lumpur in 2002, was a social gaming site and MySpace, started in 2004 and the largest platform anywhere until 2010, were the best-known players. Facebook began, as Zuckerberg relishes relating, in his Harvard dorm room and quickly grew to be the mega-corporation that it is today. Critical to its success were the invention of Facebook "friends" from whose data assumptions can be made about whole groups and population segments with similar characteristics. Similarly, the "like" button innovation that enabled users to approve and rate others' contributions, to engage in impression management and identity construction, and, crucially, also permitted Facebook to track users as they move from site to site, thus accruing more and more data.

Thus Facebook "connects" users with other acquaintances, family members, groups and so on, as heavily advertised from the beginning. But it also connects users with unseen others – the data brokers, developers, advertisers, political campaigners and snake-oil vendors that pay Facebook for data about these valuable connections. This is Facebook's business model, which falls squarely into the surveillance capitalism category. People are attracted to the site and encouraged to spend more and more time there so that their attention, their interests, the details of their daily lives, may be sold to the highest bidders. As data are donated, unwittingly, or at least only vaguely perceived, by users, so the data are used to profile those users and their friends and acquaintances, including those with no Facebook account. As with all social media, these interactions with the site are the source of value. And their aim is not merely to predict but also to shape lives and lifestyles.

The idea of connecting people sounds innocuous and attractive. Two billion users and more rely on Facebook for a host of connections and it clearly meets needs, including those needs engineered by Facebook's psychologists. However, there have been strong indications, from the outset, that Facebook's aims were

not limited to their lofty social aspirations to connect users "with those whom we love." As one player in the current scandal, Sandy Parakilas, a former data manager for Facebook, put it when asked about the privacy concerns of users: Facebook "prioritizes the growth of users, the growth of the data they can collect and their ability to monetize that through advertising . . . [;] . . . those . . . are the metrics that the stock market cares about" (Stahl 2018). To understand surveillance capitalism better, however, we must turn to Shoshana Zuboff.

Zuboff holds an important place in Surveillance Studies, for her 1988 analysis, *In the Age of the Smart Machine*. Her brilliant finding was not just that automation allows machines to lighten the load on labour and to capture and develop within software skills previously perfected by human beings but that the deep difference lies in the ways that automation also informates. The application of information technology makes the tasks more transparent to managers who can use their enhanced knowledge to control more precisely the way work is done. The workplace thus becomes more intensively surveillant.

Zuboff's new book, *The Age of Surveillance Capitalism* (2019) describes in more detail the emergence of "surveillance capitalism" that builds on but goes far beyond the argument of "Smart Machine." Here is doggedly persistent social research at its best. Animated by Google's business model, as found in the work of Hal Varian (Google chief economist), it is the source of Google's massive value of over $600B (Apple is $750B+; Microsoft: $521B; Amazon $433B; Facebook: $420B). Zuboff's work depends on extensive interviews with all the key leaders of the big five internet corporations – Google, Facebook, Amazon, Microsoft and Apple – and produces a very largescale analysis of the phenomenon that is breathtaking in its scope and boldness. Surveillance capitalism evidences several key features.

Google's secret of profitability is what Zuboff describes as "unilateral surveillance and behaviour modification" (Zuboff 2016). It sells real-time access to everyday life, aiming to change behaviours at scale, through data capture, analysis, and reward/punishment. The logic of accumulation determines what is measured or ignored and how resources are allocated. Computer mediation "now means that the world is visible, knowable and shareable in a new way" (Zuboff 2015, 76) It is corporations that gain access to everyday life and today this exempts almost no one in societies dependent on digital infrastructures. There is no transaction with users or consumers, however. Straight extraction is all that occurs at that level. The trade in data is entirely between large corporations.

Going beyond some other authors (e.g. Boltanski and Chiapello 2018), Zuboff concludes that this adds up to a new logic of capital accumulation, far beyond old supply and demand approaches that, she argues, up until the recent past tied capitalism more or less to population needs. It is characterized, she avers, by a combination of digital dependence, indifference and neoliberalism. It uses prediction to eliminate uncertainty, that may produce anxiety if not other emotions. But it also undermines social trust, cohesion, familial bonding, and binding contracts and promises. As she warns, it finally severs those already frail and frayed relationships between capitalist corporations and their employees, their consumers and their users.

The key dimensions of surveillance capitalism are as follows: (i) multiple data sources are exploited, pervasively recording everything. StreetView, is a key example, which has, of course led to protest and legal action in several countries. Another is Sidewalk Labs, currently making a bid to revitalize a whole area of derelict wharves on the Toronto waterfront, reviving the district "from the internet up" as Alphabet, for Google, puts it (*Economist* 2018); (ii) data extraction occurs, a one-way process, lacking relationship or structural responsibilities yet dependent on "signals of subjectivity" (Zuboff 2015, 79). Extraction sums it up. Ordinary internet users have no say. Data are expropriated without permission (unless one counts "terms of use") or apology. The "formal indifference" to users and consumers is visible right here, and to employees, in physical plants such as Amazon's in Seattle where making impossible demands on workers is how managers themselves define their task (Kantor and Streitfeld 2015); and, (iii) analytics means authority (spiritual) is supplanted by technique (material), producing "anticipatory conformity."

Google was one of the first to use analytics to increase the relevance of ads to users but also, crucially, to repurpose the growing cache of behavioural data especially after the advent of social media. The market exchange, as noted earlier, is not with those users but with other corporations. The term "data exhaust" downplays the reality of what is being captured from users but in reality, argues Zuboff, it is "behavioural surplus" (Zuboff 2016) For her, this mirrors geographer David Harvey's argument concerning "accumulation by dispossession" (Harvey 2004). In so doing, even rights are erased, creating basic threats to both dignity and democracy.

With such a dramatically successful development within capitalism it is all too easy to succumb to complacency or cynicism – what can be done in the face of such accelerating new logics of accumulation and data dependence? The corporations involved are indeed the highest valued on the planet and the impunity with which they operate is staggering. These factors should not be minimized. Yet to ask only such questions is to ignore those who experience surveillance in everyday life and whose responses are far from monochrome. "Anticipatory conformity" may well express part of the response and certainly, much of the everyday world of surveillance experience shadows surveillance capitalism, such that doing surveillance on others using social media, or on oneself, through self-tracking using wearables, does occur.

Of course, as a number of analysts has argued, the dominant world of surveillance capitalism pulls many into its seductive force-field. And, equally true, we have all been exposed, for decades, to the alluring sirens of consumerism, now in digital dress. The very concept of freedom is pitifully reduced to individualistic self-determination and even to consumer choice that even extends beyond mere purchasing. And, more specifically, dominant forms of big data surveillance are echoed in the dominant aspects of surveillance culture. As I say, few seem to question what is happening and, apparently, little or no resistance is offered to the secondary uses of people's "behavioural surplus," the mundane, everyday data that our machines exude constantly.

But there is also evidence, not only of dominant influences but also of residual and emergent approaches (Williams 1977) in which older outlooks guide surveillance imaginaries and practices, or newer ones offer forms of querying or resistance to surveillance. Much work has been done on North America and Europe but studies of the Global South are also appearing, demonstrating that surveillance capitalism is expanding its frontiers of accumulation. Also, different age cohorts are represented here, but the two may also join forces – for example in privacy-promoting or digital activist groups such as those seeking open access, such as OpenMedia in Canada. But to understand these, we have to consider another important concept, surveillance culture.

Surveillance culture

When I first worked in surveillance studies, more than 30 years ago, the key issues were government – the "surveillance state" – policing and workplace surveillance, often crystallized in the iconic video camera. Computerization was well under way and this affected each of the three areas, plus also in the use of credit cards, which began in the 1960s. Surveillance practices became more prominent in the area of consumption but were experienced in tangible, paper-based forms such as the rise of junk mail that targeted more and more specific groups of consumers. Surveillance was spilling over the rims of its previous containers and talk of "surveillance society" began in the mid-1980s and took hold by the turn of the century as these practices became more pervasive.

But by the first decade of the twenty-first century things were changing again. The new technologies – seen especially in so-called dot-com companies, were faring badly. New opportunities appeared following 9/11 as security industries went into a higher gear and soon afterwards as social networking developed from MySpace and Friendster into the "social media" that are commonplace today. These are reflected in other activities, subtly at first, in which ordinary citizens were invited to "say something" about perceived security breaches, to report "unusual" events or objects, or to "tip off" authorities on the one hand, and on the other, as ordinary users of social media began to exploit the new possibilities, to check up on each other in more direct ways and even to conduct private investigations into strangers' lives. A new "culture of surveillance" was taking shape.

There are many ways of considering surveillance culture. On the one hand, it has to do with the experience of surveillance in everyday life, as people negotiate ubiquitous cameras in public and private spaces, pass through security areas such as those at airports, encounter embedded surveillance in buildings, vehicles and proliferating devices, each of which collects, stores, transmits, analyzes and acts on data. And on the other hand, surveillance culture exists where people play a more active role. This may be changing personal practices in "watched" public spaces such as streets, malls or airports or in new modes of checking up on the lives of others known and unknown using conventional search engines or more likely through social media.

All these aspects of surveillance culture, whether the experiences of surveillance or engagement with surveillance take their place within everyday surveillance imaginaries and practices. The former are the way that actors see the world of surveillance and their part in it, which includes a sense of how things should be and sets of warning bells when something seems not quite right. We may expect, for instance, to be subject to have our bags checked at the airport and perhaps to enter a biometric or have our hands swabbed. Equally, we anticipate that certain online sites will require agreement with terms of service, or that the door will remain locked if our entry card is not up-to-date or that the car will not start if our blood alcohol level is too high.

Mundane surveillance practices work with this, such that people learn how best to get through security without delay – people who think they may be thought of as Middle Eastern or Muslim will plan this well in advance – to click acquiescence with the terms of service regardless of whether or not they were actually read, or to follow someone else through the security door rather than use the appropriate entry card. In the world of social media, people are most likely to check up on others known to them although about a third of American, British or Canadian citizens will check up on strangers, despite the fact that such snoopers believe that these people would be annoyed, upset or embarrassed if they knew (Smith and Lyon 2013).

Of course, surveillance culture is volatile, complicated, as a leading analyst, danah boyd (2015), observes. To examine surveillance imaginaries, by which people envision the world of surveillance and see their place within it, and their surveillance practices, which is how they engage with surveillance, is to find a wealth of evidence of heterogeneity, not mere homogeneity.

For a start, while a majority (in the US) have cellphones, around a quarter do not yet have a smartphone (Pew Research Centre 2018). Also, context is crucial. Much evidence shows that knowledge of surveillance is widespread but that only certain kinds of data gathering may be viewed negatively. I may share my healthcare information with my doctor and some family members, but be much more circumspect if I think that an insurance or pharmaceutical company may seek to see it. This is not just a matter of caring more about watchers who are known (friends, family) than unknown (corporate marketers). Many judge surveillance practices (not necessarily recognizing them as surveillant) according to criteria not of "privacy" but of "fairness" (Kennedy et al. 2015) and act accordingly.

For a long time, much research has demonstrated that surveillance involves gathering data to enable populations to be categorized so that different groups can be treated differently. This is social sorting (Lyon 2003). Dominant forms, especially those using so-called big data, tend to reinforce already existing disadvantage and marginalization. A fine example is *Automating Inequality* by Virginia Eubanks (2018). And as a recent *Data & Society* report concluded, "Marginal populations may be subjected to increased surveillance by both public and private actors. If predictive algorithms deem them to be "at-risk," they may be labeled as such and further marginalized" (*Data & Society* 2014).

The exclusionary impetus of surveillance as social sorting is thus augmented among those who are already vulnerable in racialized, gendered, class-based and other categories. But even this does not mean that the outcomes may be taken for granted. While vulnerable populations may find their life-chances further restricted by big data practices, at the same time, "gaps in data . . . might be used to empower or assist groups rendered invisible by targeted data collection." Women on welfare will subvert surveillance in order to look after their children (Gilliom 2001); brown-skinned air-travellers will perform for security in risk-reducing ways (Akseer 2017); poorer people in the housing market will use real estate classifications to mitigate their position (Burrows and Gane 2006); and smart phone users will exchange SIM cards for different purposes. Surveillance practices are manifold.

Situating surveillance culture, surveillance capitalism

To speak of surveillance culture and surveillance capitalism is clearly to engage in large-scale social analysis relating to twenty-first century surveillance. To begin, surveillance culture requires that common definitions of surveillance have to be stretched. Conventional definitions often start from an "operator" perspective that sees surveillance as something that happens to social actors, whether negotiating airport security, walking down the street under the gaze of cameras, or becoming aware that using Instagram and WhatsApp on those smartphones means that personal identity, preferences and whereabouts are known. But in fact, as I argue here, surveillance is also something that people now engage with in daily life, sending images of incidents to the police, installing home security systems or checking up on others, including strangers, using Facebook or some other social media platform. In everyday life, ordinary people contribute to a growing culture of surveillance; watching is becoming a way of life.

In each case just mentioned, everyday surveillance is facilitated by relatively new technologies that have proliferated in recent decades, even as surveillance has risen in cultural significance due to its prominent use in government, corporate and security contexts. Surveillance cameras, for example, were encountered increasingly in urban areas from the 1970s but especially from the 1990s and thus became part of quotidian experience. For various reasons, CCTV systems not only came to be viewed as viable means of combating crime and disorder and even of providing safety on the street but were also marketed for domestic protection. In some Brazilian cities, for instance, the issue is fewer public cameras "intruding" on private spaces than privately owned and operated cameras with capacities to watch public spaces (Firmino 2018). Thus, the slide from rarely encountering to routinely experiencing and to regularly engaging with surveillance cameras occurred. Surveillance is thus normalized and taken for granted, even though debates persist at every level about its appropriateness and efficacy.

But these developments in surveillance culture occur in the same world characterized by the political economic realities already discussed; "surveillance capitalism," in which contemporary data extraction is profoundly implicated.

That is, for example, as users go online to use Google, Facebook or even – as academics – ResearchGate (Lyon and Melgaço 2019), scraps of data are sucked up as a vacuum cleaner sucks away detritus from rugs or sofas. But this digital dust does not go to landfills. Someone – Google was first! – saw value in it and now it is monetized to make millions. In other words, the surveillance culture has an intimate and mutually-informing relationship with surveillance capitalism. What is that relationship? As I hinted earlier, the dominant aspects of surveillance culture often play into surveillance capitalism, facilitating and normalizing it. And by the same token, much of surveillance culture depends on and is nurtured by surveillance capitalism.

However, it is important to observe that focusing only on the operator aspects – the vacuum cleaning – or on the complacent and compliant aspects of surveillance culture can easily produce a sense of hopelessness or at least, cynicism. The "operators" will insist that the technological changes shaping the digital era are really unstoppable and that not to be data-driven is to miss out on efficiency and profitability or at least that law and regulation will never "catch up." And those comfortable with surveillance culture will say that its convenience and efficiency in making desired connections with others is worth any minor quibbles about things like privacy or civil liberties.

To counteract the sense that nothing can be done I note that analyses of everyday surveillance imaginaries and practices (Taylor 2003) – "surveillance culture" – indicate that in fact a variety of responses is possible and, increasingly, visible. True, some shrug off the sucking up of data as something inconsequential; who cares? But beyond such dominant modes are residual approaches that question users donating data with no apparent return (Andrejevic 2013), and emergent modes that try to resist by arguing for new forms of regulation or by using technical means in digital judo moves (Dupont 2008). In the mid-twentieth century, some early studies of TV feared the growth of propaganda and the negative impact of the new medium, that "cultivated" viewers. But more subtle studies showed just how much "critical viewers" also exist, reading the news or interpreting the shows in myriad ways. Arguably, something similar is happening now, in relation to the digital and to surveillance. It is consonant with Engin Isin and Evelyn Ruppert's observation, that while some internet users see themselves as simply subject to power, others believe that they can make a difference. They are subjects of power (Isin and Ruppert 2015).

What these authors note, in *Being Digital Citizens*, is this. While "subjects" is a useful and illuminating word, it has to be thought of in two ways at the same time. People are subjects to power in that everyday lives are profoundly affected positively and negatively by data and the internet – particularly, as argued here, by surveillance capitalism. But simultaneously people are subjects of power in that they may demonstrate subversive as well as submissive behaviours in online life – or "onlife" (Floridi 2015). Digital citizens come into being, in part, as data politics begin to form themselves in recognizable ways. Our very relationship as citizens in digitally dependent societies is now mediated by the internet and by data. And

as "digital citizens" make rights claims about those data, they do so prompted and provoked into self-governing and by attempting to exert political influence through such claims. Thus, while today's strategies of power are already being emulated by others in subordinate positions they are also evaded, questioned and subverted by everyday tactics.

Data politics and an optics of hope

Having tried to make the case for seeing together the two phenomena of surveillance capitalism and surveillance culture, I turn towards a more normative conclusion. I want to press the foregoing argument further. For a meaningful data politics to emerge, it would seem that human dignity and especially agency need not only to be seen in diverse responses to surveillance capitalism but also to be reasserted and encouraged. Given the manifest disrespect for dignity and fairness suffered under surveillance capitalism, indignation should rightly feature as an aspect of such social analysis. Surveillance capitalism and surveillance culture cannot simply be studied dispassionately, however carefully and accurately the social data are presented. They affect identities, how people see and present themselves (subjectivities), life chances, the ways that opportunities open or close depending on the consumer and other categories in which, today, everyone is placed (social sorting) and democratic participation, or how far we can vote or whether or not that vote makes a difference (politics). In other words, human life is in many ways – put concisely, in relation to subjectivities, social sorting and struggle – negatively affected by surveillance capitalism.

Indignation is necessary but not sufficient, however. Analytically, the work of Michel de Certeau (1984) opens doors to a more hopeful sociology of surveillance for the twenty-first century. While he freely acknowledges the strategies of power visible in late modern consumer capitalism, he urges that attention also be focused on the tactics visible in everyday life – tactics that do not simply mirror the dominant political economy, in our case, of surveillance capitalism. In the present context, this means paying careful attention to surveillance imaginaries and practices. These include the repurposing of technologies beyond what was envisioned by their designers, resisting social media marketing and finding fresh ways of using media – for instance, by accenting quality, not quantity. For example, Catchpool – whose tagline is "catch the best, leave the rest" – reduces mere noise for the sake of valuable forms of sharing, or Mighty Networks produces and distributes customizable tools for people to create their own networks (Laurenson 2016).

As so often, reminders are needed that de Certeau's tactics may operate at many levels. There are the everyday online interactions in which micro-responses to surveillance capitalism – even though it may not be named as such – occur and which, when magnified by social media, could make a difference. But there are also more deliberate activities that turn simple responses into rights claims and that, again, could be amplified by shared activities, this time through some of the many rights-claiming groups now springing up to alert users to possible abuses and

remedies. These are often associated with groups that also have technical expertise to assist those with none, to clarify their claims and their targets. The tactics may also be associated with more conventional and formal rights claims, made in relation to privacy and data protection legislation. All these need to be filled out with in-depth research, but of their existence there is little doubt. And moments such as the Facebook-Cambridge Analytica scandal, as the moment of the Snowden disclosures or the post-9/11 security surveillance overreach, are critical ones for the discovery that what seemed to be personal problems are shared public issues.

Here, I am not so much prescribing ways forward as proposing some open questions about where surveillance culture is heading and how some emerging trends might channel it in fresh ways. Contemporary cultural developments may foster human flourishing. Even a recent British report on data management couches its aims as the pursuit of human flourishing (Royal Society 2017). Along with a quest for fairness – for instance in the Cardiff University research group dedicated to "data justice" – both would-be responsible consumers and wider human rights groups assert more human scale approaches against tendencies towards a colder and more calculating surveillance capitalism.

This is where I believe that the notion of utopia-as-method (Levitas 2013) is instructive. This is not the familiar fictional accounts of idealized worlds, but rather, critical accounts of current cultural directions. These also act as potential means of proposing and promoting alternative futures that embody holistic, reflexive and democratic imaginaries and practices. In other words, utopia-as-method contributes to the common good as shared values – chosen wisely, by conviction and conscience, not consumer criteria – especially in relation to rights and responsibilities. One area that such shared understandings develop is in the realm of popular cultural forms, including utopias and dystopias.

The commonest metaphors in surveillance culture today still come from Big Brother, now largely rendered obsolete – in its details, not in its humane thrust – by the rise of surveillance capitalism. But this does not mean either that Orwell is irrelevant or that other metaphors and memes are unavailable. They may be found, for instance, in some older fiction such as *Lord of the Rings* or in the contemporary utopian/dystopian fiction of *The Circle* and of TV series such as "Black Mirror." The task of the social sciences, alongside some very well-informed contributors to literary criticism (Marks, 2015, Rosen and Santesso 2013) should try to understand how the new metaphors are mobilized as means of comprehending and acting in relation to surveillance.

Surveillance capitalism is a newly dominant social, economic and political formation. But to understand surveillance only in those terms is to see it from an exclusively operator perspective. Raising awareness about its actual mode of operation and its erosion of relationships and rights is a vitally worthwhile task. However, also considering possible tactics that might destabilize or deflect some of its consequences will set the tone for a struggle that is already under way. It is unclear how the 2018 Facebook debacle will play out. However, the very fact of raised consciousness and widespread, publicly discussed uncertainty

about the balance between the pros and cons of involvement with Facebook is a sign that data hegemony is far from complete, and of the ongoing volatility of surveillance capitalism.

Seeking workable alternatives as well as promoting limits to the expectations of 24/7 access – the "always-on" phenomenon – alongside the increasing pressure to find acceptable modes of regulating social media could mean that a turning point is being reached. Looking at the everyday life of surveillance as it is experienced, imagined and practiced – surveillance culture – offers not just a complementary and necessary perspective than the rather too prevalent paranoia, complicity and defeatism associated with the critique of social media and surveillance capitalism, but the potential for an optics of hope. Why? Because the symbiotic growth of surveillance capitalism and culture will only be interrupted if the latter becomes more conscious of itself and more willing to ask basic questions: Do we really need this? How does it contribute to the common good and human flourishing? If those cultural questions, relating to how we actually live our lives, generate a fresh data politics – green shoots of which are already appearing – then those hopes will begin to be realized.

Note

1 Earlier versions of this chapter were presented at ASA Montreal August 2017; Universidade Federal do Rio de Janeiro October 2017; CICC atelier Université de Montréal, Nov 2017; IAS Loughborough University Nov 2017. Brief version for the Saturday Club, Kingston, May 2 2018.

References

Akseer, Tabasum. 2017. *Understanding the Impact of Surveillance and Security Measures on Canadian Muslim Men: A Mixed Methods Approach*. PhD Dissertation, Queen's University.

Albrechtslund, Anders. 2008. Online social networking as participatory surveillance. *First Monday* 13(3). Available at: https://firstmonday.org/article/view/2142/1949

Andrejevic, Mark. 2013. *Infoglut*. London: Routledge.

Bauman, Zygmunt and David Lyon. 2013. *Liquid Surveillance*. Cambridge: Polity.

Boltanski, Luc and Chiapello, Eve. 2018 (new edition). *The New Spirit of Capitalism*. London: Verso.

boyd, danah. 2015. *It's Complicated: The Social Lives of Networked Teens*. New Haven: Yale University Press.

Burrows, Roger and Nicholas Gane. 2006. Geodemographics, software and class. *Sociology*, 40(5), 793–812.

Christl, Wolfie. 2017. Corporate Surveillance in Everyday Life: How Companies Collect, Combine, Analyze Trade and Use Personal Data on Billions. A Report by Cracked Labs, Vienna, June. Available at: http://crackedlabs.org/dl/CrackedLabs_Christl_CorporateSurveillance.pdf

Data & Society. 2014. *Data & Fairness*. Available at: https://datasociety.net/initiatives/data-fairness/

Davies, Harry. 2015. Ted Cruz using the firm that harvested data on millions of unwitting Facebook users. *The Guardian*. December 11. At www.theguardian.com/us-news/2015/dec/11/senator-ted-cruz-president-campaign-facebook-user-data

de Certeau, Michel. 1984. *The Practice of Everyday Life*. Berkeley: University of California Press.

Dupont, Benoît. 2008. Hacking the panopticon: Distributed online surveillance and resistance. *Sociology of Crime, Law and Deviance* 10: 259–280.

Economist. 2018. The world's first neighbourhood built "from the internet up." *The Economist*. At www.economist.com/business/2018/05/03/the-worlds-first-neighbourhood-built-from-the-internet-up

Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*. New York: St Martin's Press.

Firmino, Rodrigo. 2018. Smartphones Smart Spaces? O uso de mídias locativas no espaço urbano em Curitiba, Brasil. *Eure* 44(133): 253–273, September. At www.researchgate.net/publication/326590280_Smartphones_Smart_Spaces_O_uso_de_midias_locativas_no_espaco_urbano_em_Curitiba_Brasil

Floridi, Lucien. Ed. 2015. The Onlife Manifesto: Being Human in a Hyperconnected Era. SpringerLink. *At* https://www.springer.com/gp/book/9783319040929

Gilliom, John. 2001. *Overseers of the Poor: Surveillance Resistance and the Limits of Privacy*. Chicago: University of Chicago Press.

Harvey, David. 2004. The 'new imperialism': Accumulation by dispossession. *The Socialist Register* 40. At https://socialistregister.com/index.php/srv/article/view/5811/2707

Isin, Engin and Evelyn Ruppert. 2015. *Being Digital Citizens*. New York: Rowman & Littlefield.

Kantor, Jodi and David Streitfeld, 2015. Inside Amazon: Wrestling big ideas in a bruising workplace. *The New York Times*. August 15. At www.nytimes.com/2015/08/16/technology/inside-amazon-wrestling-big-ideas-in-a-bruising-workplace.html

Kennedy, Helen, Dag Elgesem and Cristina Miguel. 2015. On fairness: user perspectives on social media data mining. *Convergence* 23(3), 1–19.

Laurenson, Lydia. 2016. Social media platforms can be built around quality, not scale. *Harvard Business Review*. Available at: https://hbr.org/2017/04/imagining-a-social-media-platform-built-around-quality-not-scale

Levitas, Ruth. 2013. *Utopia as Method: The Imaginary Reconstitution of Society*. London: Palgrave Macmillan.

Lyon, David. Ed. 2003. *Surveillance as Social Sorting: Privacy, Risk and Digital Discrimination*. London: Routledge.

Lyon, David. 2017. Surveillance Culture: Exposure, Engagement and Ethics in Digital Modernity. *International Journal of Communication* 11: 824–842.

Lyon, David. 2018. *The Culture of Surveillance: Watching as a Way of Life*. Cambridge: Polity.

Lyon, David and Lucas Melgaço. Forthcoming. Surveillance and the quantified scholar. *International Studies Perepectives*.

Marks, Peter. 2015. *Imagining Surveillance*. Edinburgh: Edinburgh University Press.

Marwick, Alice. 2012. The public domain: surveillance in everyday life. *Surveillance & Society* 9(4): 378–393.

Pew Research Centre. 2018. Mobile Fact Sheet. February 5, Available at: www.pewinternet.org/fact-sheet/mobile/

Rosen, David and Aaron Santesso. 2013. *The Watchman in Pieces*. New Haven: Yale University Press.

Royal Society. 2017. Report on Big Data at www.data-management-governance.pdf

Smith, Emily and David Lyon. 2013. Comparison of Survey Findings from Canada and the US on Surveillance and Privacy from 2006 and 2012. *Surveillance & Society*, 11(1/2). At https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/survey

Stahl, Lesley. 2018. CBS News: Aleksandr Kogan, the link between Cambridge Analytica and Facebook, www.cbsnews.com/news/aleksandr-kogan-the-link-between-cambridge-analytica-and-facebook/April

Taylor, Charles. 2003. *Modern Social Imaginaries*. Durham: Duke University Press.

Trottier, Daniel. 2017. Digital vigilantism and the weaponization of visibility, *Philosophy and Technology*, 30 (1): 55–72.

Williams, Raymond. 1977. *Marxism and Literature*. Oxford: Oxford University Press.

Zuboff, Shoshana. 1988. *The Age of The Smart Machine: The Future of Work and Power*. New York: Basic Books.

Zuboff, Shoshana. 2015. Big Other: Surveillance capitalism and the prospects of an information civilization, *Journal of Information Technology* 30: 75–89.

Zuboff, Shoshana. 2016. The secrets of surveillance capitalism. *Frankfurter Allgemeine*. 5 March. At www.faz.net/aktuell/feuilleton/debatten/the-digital-debate/shoshana-zuboff-secrets-of-surveillance-capitalism-14103616.html

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism*. New York: Public Affairs.

PART II
Worlds

5

MUTUAL ENTANGLEMENT AND COMPLEX SOVEREIGNTY IN CYBERSPACE

Ronald J. Deibert and Louis W. Pauly

When the Internet first emerged, many predicted that it would present a major challenge to the power of states in general and to the effective control of authoritarian states in particular (Johnson and Post 1996). More recent commentary has emphasized the opposite: that the Internet expands and intensifies the capacities of states within and across conventional territorial boundaries. While we now see clearly how social media and other digital technologies have empowered non-state actors in civil society, the verdict is quite mixed as to their ultimate impact on the sovereign authority of the state.[1]

Domestic-level information controls are today reinforced by norms promoted by states like China and Russia that try to shift governance away from multi-stakeholder or pluralist models toward more state-centric approaches. Even liberal democratic countries have lately been moving in the direction of territorially-defined policies of cyberspace governance through laws aimed at data localization and through the establishment of "cyber commands". While efforts to re-territorialize cyberspace are undeniable, the extent to which states depend on mutual restraint to project power in and through cyberspace has been obscured. Extraterritorial projections of state power in this sphere are expanding, deepening, and becoming more elaborate. The most extensive of these projections come from the United States, but even the most autocratic regimes associated with efforts to promote "Internet sovereignty" today rely on the openness of cyberspace.

States are exercising extraterritorial power to acquire data about the world around them: to anticipate, analyze, and interdict threats; to shape the strategic environment to their advantage; to promote their interests via the movement of goods and services, information, and capital. They are also using new communication technologies to broaden military command and control systems. The combined if not fully intended "network effect" of such extensive projections of power in and through cyberspace is to frustrate individual strategies aimed at territorial insulation.

This effect today is well-described as mutual entanglement (Nye 2017). The capacity of states to project power domestically and extraterritorially rests on the material opportunities opened up by cyberspace itself, and that openness thwarts efforts to build impenetrable border controls. As states aim to shape cyberspace to their strategic advantage, their governance domains both expand and contract. Specific policies are continually being reconfigured in a dynamic if not necessarily symmetrical context of interaction. The legitimacy of those policies may be contested, but they rest on an apparently adequate degree of acquiescence internally and externally.

Drawing from recent research into state espionage and targeted digital attacks, as well as evidence now in the public domain from the Edward Snowden disclosures that cannot be ignored, this chapter provides an overview of extraterritorial projections of state power in and through cyberspace, from the United States to cases involving highly opaque autocratic regimes. This evidence suggests that efforts to bring digital networks back under territorial control are undercut by operations designed to use those networks for domestic surveillance and external security. In the end, re-territorialization strategies in cyberspace are self-limiting. The chapter concludes by sketching implications for sovereign authority in a dynamic system. The mutual entanglement characteristic of cyberspace today profoundly complicates state strategies aimed at either anarchical fragmentation (where no one sets governing rules) or unquestioned hegemony (where rules are set by a dominant power) (Deudney 2007; Ruggie 1993).

The territorialization impulse in cyberspace

The OpenNet Initiative (ONI) – a university-based research project using a mixed methods approach to documenting Internet censorship – has conceptualized state power over cyberspace within territorial boundaries in "generational" terms (Deibert 2015, 2017; Deibert and Rohozinski 2008). First generation controls refer to defensive Internet censorship systems erected at national borders, with governments restricting their citizens' access to online resources, the Great Firewall of China being the archetypal example. Internet filtering typically involves special software or hardware placed at key network chokepoints that inspect requests for web content, blocking those that are restricted from reaching their destinations. ONI tested for national-level Internet filtering in more than 70 countries and found evidence in more than 45 (Deibert et al. 2008; 2010; 2012). The number is likely expanding quickly, since many countries have begun censoring content involving the sexual exploitation of children, hate speech, and terrorist threats.

Second generation controls refer to government measures to control cyberspace domestically through laws, policies, and other sorts of Internet policing, often undertaken with the cooperation, coercion, or co-optation of private companies. Examples include content removal requests, compelled access to customer data, and the application of defamation or libel laws to Internet content. Sometimes second-generation controls are applied secretly, making documentation challenging for researchers. Occasionally we see glimpses of these controls through the window of

private sector transparency reports, such as those published by Google, Microsoft, or Twitter. The remarkable Vodafone Law Enforcement Disclosure Report, for example, extensively documented country-by-country requests for customer data (Vodafone 2014). Researchers have also employed reverse engineering methods to uncover hidden surveillance or censorship functions built inside popular applications, such as the surveillance embedded inside the Chinese version of Skype (Dalek et al. 2015; Knockel, McKune and Senft 2016; Knockel, Senft and Deibert 2016; Villeneuve 2008). It is accurate to say that second-generation controls have become more complex, penetrating deeper into civil societies and filtering communications through a thicket of rules, laws, and practices.

Third generation controls refer to the use by states of more "offensive" methods, such as targeted surveillance, digital espionage, and disinformation campaigns. If first generation controls sought to bolster borders, and second generation controls deepened the internal reach of state agencies, third generation controls are projected outwards. Although varying in resources and capabilities, many governments' armed forces and intelligence agencies have developed aggressive external operations. Growing demand for offensive capabilities has produced a rapidly expanding market for computer network attack and surveillance products and services developed by private companies. These firms range from Cold War giants like Raytheon and Northrop Grumman to more obscure "niche" entities, like Italy's Hacking Team, the UK's Gamma Group, or the Israeli "cyber warfare" company, the NSO Group (Harris 2014). The overall industry is growing at an annual rate of 24% per year and will likely exceed USD $600bn in annual revenue by 2023 (Stiennon 2016).

Across all three generations, cyber security has risen to the top of policy agendas, driven by repeated instances of large scale data breaches, vulnerabilities to critical infrastructure, competitive issues, and domestic political concerns (Deibert and Rohozinski 2010). To the three generations of controls, moreover, might be added a fourth: the efforts of some states to negotiate governance agreements at regional and international levels. Over the last several years, for example, a coalition of like-minded countries led by China and Russia, using the rhetoric of "Internet sovereignty" and leveraging the opportunity presented by the Snowden disclosures, has sought to move governance practices away from what they perceive as its current US-dominated system to one centered around the United Nations and organizations like the International Telecommunications Union (Deibert and Crete-Nishihata 2012).

Predictions of Internet "fragmentation" and a retreat toward "Cyber Westphalia" have become prominent (Demchak and Dombrowski 2011; Dombrowski 2016). Different sources have specifically been identified: filtering and blocking websites, social networks or other resources offering undesired contents; attacks on such networks and resources; digital protectionism blocking users' access to and use of key platforms and tools for electronic commerce; centralizing and terminating international interconnections; attacks on national networks and key assets; local data processing and/or retention requirements; architectural or routing changes to keep data flows within a territory; prohibitions on the transborder movement of certain categories of data; strategies to construct nationally bounded "Internet

segments"; and international frameworks to legitimize restrictive practices (Drake, Cerf and Kleinwächter 2016).

Of these, the so-called "data localization" trend accelerated by specific reactions to the Edward Snowden disclosures is worth special emphasis. Those disclosures revealed intensive electronic intelligence-gathering by the US National Security Agency (NSA) and its close allies. In reaction, many others began insisting on the holding of local data centres inside national jurisdictions and tightly restricting trans-border processing for certain classes of data. Whether such restrictions can actually prevent effective surveillance or, in the case of official investigations, reduce reliance on cumbersome mutual legal assistance treaties (MLAT), is questionable. They also raise obvious concerns that their true intent may be more domestic in nature.

In hindsight, given the externalities around Internet communications (now used by well over three billion people on a daily basis worldwide), the impulse behind expanding state control efforts was foreseeable. It now seems inevitable that states would be ever more focused on trying to shape information environments quickly becoming integral to all aspects of society, from the cultural to the economic and political. Among other things, high-profile terrorist acts certainly encouraged citizens to demand such efforts. As obvious as such an impulse may now seem, however, its implications should not be exaggerated. It constitutes only one dimension of a complex process involving the extraterritorial projection of power by other states in and through cyberspace itself. The next sections surveys recent research illustrative of that process and its consequences.

The United States and the transformation of cyberspace

The contemporary cyber-security policies and practices of the United States offer the clearest example of extraterritorial power projection. The American defense of a borderless, open internet may simply be depicted as based entirely on liberal values and ideals, and conveniently contrasted with "territorializing" processes of states that oppose this agenda. The US posture is actually more complicated. Its "Internet freedom" agenda is arguably more a function of interests than values. It is in many ways a discursive or ideological support for the projection of US power in global cyberspace. In this respect, it is analogous to the US position on treating the oceans and outer space as a "commons". The free movement of information globally (just as with free navigation of navies and satellites, and to a lesser degree, aircraft) serves global hegemonic power, not because US policymakers believe in the ideal of the open commons (although some very well might) but because sustaining a position of dominance depends on the ability to move goods, services, information, and capabilities across cyberspace.

US power projection is also connected both to long-term interests and to a changing threat environment. The US now operates nearly 800 military bases in more than 70 countries and territories worldwide (Vine 2015). This extended footprint is woven together by a bristling infrastructure of digital communications, today including 131 government and 149 military satellites in orbit as well

as another 273 US-owned commercial satellites (Union of Concerned Scientists 2016). The Pentagon alone operates around 7,000 unmanned aerial drones (Friends Committee on National Legislation 2015). Today a Hellfire missile strike from a US Predator drone is guided by earth-orbiting global positioning satellites (GPS) to within a few metres of its target. The missile is typically fired by an operator based in a hangar in the mainland US, working on computer screens onto which are projected high-resolution images beamed back instantly by advanced imaging sensors.

Such technological advances, of course, track the emergence of the United States as a global superpower. They coincide with the development of the Internet itself and by earlier innovations in telecommunications, including undersea cables and digital computing systems now global in scope (Starosielski 2015). The United States, in fact, enjoys a distinct "home field advantage" with respect to much of the geopolitics of cyberspace. Most of the Tier 1 telecommunications companies that operate the backbone of the earth's communications systems are headquartered in the United States; the largest software, social media, device, and Internet service providers are still mostly American (Deibert 2012). As a result, many firms can be compelled or quietly enlisted into US government policing and intelligence efforts – a lesson not lost on other governments. One of the more interesting consequences of the Snowden disclosures, however, has been the rolling out of consumer level end-to-end encryption by US-based companies. The consequence is to deepen and extend global networks and frustrate policies aimed at strict data localization.

US intelligence agencies have long reached directly into networks physically based outside their territorial jurisdiction. Officials and their helpers can penetrate or exploit vulnerabilities at critical nodes in the global flow of communications through remote access to cables, servers, routers, wireless networks, and Internet Exchange Points (IXPs). In this regard, switches and other hardware shipped overseas are hardly invulnerable, and encryption standards through international standard setting bodies are malleable.

Consider just one very important NSA program, codenamed XKEYSCORE. XKEYSCORE provides a portal for analysts into the massive amounts of digital electronic communication data that are vacuumed up from access points around the world. The Snowden disclosures indicated that as of the late 2000s, XKEYSCORE-accessible communications data included not only emails, chats and web-browsing traffic, but also pictures, documents, voice calls, webcam photos, web searches, advertising analytics traffic, social media traffic, botnet traffic, logged keystrokes, computer network exploitation (CNE) targeting, intercepted username and password pairs, file uploads to online services, Skype sessions and more (Marquis-Boire, Greenwald, and Lee 2015). At that time, XKEYSCORE involved at least 700 servers in 150 field sites across a wide array of countries.

Observers commonly note that such programs suggest only that the United States is an exceptional power. That ignores, however, the experience of all "arms races" in history. US innovations in signals intelligence (SIGINT) practices are closely followed by its key allies. Second and third tier partners may be expected

to emulate them, and eventually so too will competitors. The Snowden disclosures may have accelerated this process, providing a "blueprint" of elite SIGINT techniques that others surely now strive to imitate.

It is important to recognize that US power projection in and through cyberspace is already partially coordinated through a long-standing and deeply institutionalized alliance system, most commonly referred to as the "Five Eyes", a partnership among the SIGINT agencies of the United Kingdom, Canada, New Zealand, and Australia. While Anglo-American history and culture, certain common political institutions and governing practices, and the experience of the Second World War still underpin the alliance, geography also accounts for much of its continuing vitality (Katzenstein 2012). The five agencies extensively exchange intelligence that ensures seamless coverage over the vast majority of international signals and telecommunications traffic. The United Kingdom alone remains a major hub for global flows of information to and from Europe, the Middle East, Asia, and the United States (Müller-Maguhn et al. 2014). Undersea cables terminate on its southwestern and eastern shores, while other linkages, including longstanding financial networks, connect it to the rest of the world. Canada (historically focused on North America and the Arctic) and Australia and New Zealand (focused on Asia-Pacific) provide their own regional complements. In recent decades, prompted by terrorist threats, other security concerns, and common economic interests, the Five Eyes have intensified collaboration with concentric rings of other states, including Denmark, France, Netherlands, Norway and then Germany, Belgium, Italy, Spain, and Sweden. In terms of actual practices of deep intelligence-sharing sufficient to construct what international relations scholars call a "security community", it makes sense today to talk about a collaborative arrangement involving at least "Fourteen Eyes".

The extraterritorial projection of autocratic power

Intense concerns about Internet fragmentation today typically center on the policies of a growing number of authoritarian regimes. Early predictions that the Internet would contribute to the demise of these forms of political rule were clearly misplaced. Autocratic governments have proven to be adept at building sweeping information control systems. Indeed, there are many characteristics of digital technologies – biometric databases, commercial spyware, and deep packet inspection systems – that can facilitate centralized rule. The publicity around the Snowden disclosures, moreover, may have accelerated moves to emulate controls pioneered by democratic states. In any event, there is no doubt that authoritarian government interference in Internet traffic – from content filtering to complete disruption of services – has become commonplace (Gunitsky 2015). Re-territorialization strategies, though, have to be viewed with skepticism.

China has commonly been seen as the progenitor for a new paradigm aimed at closing or tightly controlling cyberspace. It employs all three generations of information controls, from its Great Firewall blocking access to websites and services hosted outside of China's borders, to its extensive, legally mandated system of

social media controls imposed on domestic Internet service companies and providers. Companies employ thousands of individuals whose jobs are to censor posts on popular social media and other communications platforms. Many engineer their systems with surveillance functionalities, and all locally based companies are required to share user data with state security services upon request. Internationally, China pushes an agenda to build a new Internet governance regime assigning priority to state sovereignty and "non-interference".

Nevertheless, China has not been able to hide its own extraterritorial reach. Its external operations are most evident in vast and well documented cyber espionage campaigns, which include both global targets and an extensive transnational network of command and control servers based outside of China's jurisdiction. The US security company Mandiant (n.d.), for example, traced one of many major China-based campaigns. Known as APT1 ("Advanced Persistent Threat 1") and involving 937 Command and Control (C2) servers hosted on 849 distinct IP addresses in 13 countries, it has been convincingly linked to a unit of the People's Liberation Army.

Apart from its extraterritorial projection of power through electronic espionage campaigns, China also has extensive transnational reach through its telecommunication and software industries. Huawei, the largest telecommunications equipment manufacturer in the world, has engineered routers that by accident or design allow unauthorized access (Blue 2012). Researchers have documented, through reverse engineering techniques, massive privacy and security vulnerabilities in several China-manufactured applications, including UC Browser, QQ Browser, and Baidu Browser. UC Browser is used by 500 million people, many outside of China. Baidu Browser's software development kit, essentially a suite of code, has been adopted in tens of thousands of other applications that themselves have been downloaded hundreds of millions of times outside of China, which means that the same data collected by Baidu Browser and sent back to Baidu's servers, for possible sharing with Chinese authorities, is sent in the same way. Although there is no publicly available evidence connecting these privacy and security vulnerabilities directly to Chinese state agencies, the mere collection of such fine-grained information, coupled with well-established data retention and sharing practices inside Chinese industries, means the effect is the same. Non-Chinese national users of these applications know or should know that they are exposed to surveillance by Chinese authorities (VanderKlippe 2016). It is highly probable that China's state security organs are harvesting this information, much the same way the Five Eyes harvest information collected by western companies.

One of the more remarkable examples of China's extraterritorial projection of power in cyberspace is the "Great Cannon", a digital attack tool co-located in China's Great Firewall. It was discovered and documented by researchers at the University of Toronto's Citizen Lab in collaboration with computer scientists at UC Berkeley and Princeton University (Marczak, Weaver et al. 2015). After reports emerged of denial-of-service attacks targeting the websites of overseas critics of the Chinese government, the researchers used several network measurement techniques

to document this new attack tool. They named it the Great Cannon because it repurposes a random set of external requests for access to websites inside China and then deploys them as packets in attacks aimed at overseas websites. Functionally speaking, the Great Cannon effectively "shoots" such requests back and thereby overwhelms servers located outside of China that Chinese operators wish to silence. The very nature of the attack tool – operating at the international gateway where China's domestic networks connect with networks abroad – points to the complex ways in which territorial impulses and transnational flows of information are necessarily entangled in the contemporary practice of digital power projection.

At the same time that Chinese authorities seek vigorously to defend their Internet borders, they also are pragmatic about the need to accommodate transnational data flows, principally for economic reasons (Lindsay 2015). The Great Firewall of China therefore remains porous by intention. To cite just one example, CloudFlare, a US-based cyber-security firm recently entered into a "virtual joint venture" with Chinese web-services firm Baidu to create a unified network that makes foreign websites more easily accessible in China and allows Chinese sites to run in destinations outside the country (Mozur 2015). While the agreement may seem orthogonal to the regime's interests in strictly defending its territorial boundaries, it is perfectly congruent with the pragmatic approach the country's elites actually take to encourage economic growth. Digital networks are seen as essential in that regard, but so too are countervailing efforts to restrict the exchange of ideas that run contrary to one-party rule or that touch on taboo topics, such as religious freedom, regional autonomy, democracy, and human rights.

China's tech companies have no choice but to participate in this balancing act. The popular chat application, WeChat, provides a prominent case in point. With 806 million monthly active users, it is the most popular such application in China and the fourth largest in the world. Citizen Lab researchers undertook several controlled experiments using combinations of China, Canada, and US registered phone numbers and accounts to test for Internet censorship on WeChat's platform (Ruan et al. 2016). They found substantial censorship on WeChat but split along several dimensions. There is keyword filtering for users registered with a mainland China phone number but not for those registering with an international number. However, once a China-based user has registered with a mainland China phone number, censorship tools follow them around — even if they switch to an international phone number, or work, travel, or study abroad. In what appears to be a complete subversion of the Cyber Westphalia thesis, a company tethered to the Chinese state is projecting an Internet censorship regime far beyond China's sovereign jurisdiction.

Another vivid illustration of China's extraterritorial projection of power into cyberspace is its ambitious, though byzantine and secretive, national space program. Since it launched its first satellite in 1970, China now has 177 satellites in orbit, second only to the United States (568) and surpassing Russia (133). These satellites include those whose purpose is communications, navigation, civil defense, remote sensing and surveillance, as well as science, and environmental monitoring. China also has a manned space program and ambitions to land a man on the Moon

by 2023. One of the cornerstones of China's space program is commercial launching capabilities, much of which have implications for the future of cyberspace. Its Long March (Chang Zheng) family of rockets is responsible for 155 satellites currently in orbit, second only to the Ariane family operated by a European consortium of countries (which is responsible for 200). That the supposed archetype of Cyber Westphalia is also one of the world's leading purveyors of satellite-based global monitoring systems underscores the need for conceptual adjustment.

Iran is another country often cited as a prime mover in the fragmentation of the Internet, but its actual practices too are more complicated in nature. The country has one of the most extensive national Internet filtering systems, and its controls embody all three generations outlined earlier. In recent years, the country has created several new agencies to oversee information controls, including the Supreme Council of Cyberspace, the Cyber Army, the Committee Charged with Determining the Instances of Offensive Content, and the Cyber Defense Command. Iran has been developing plans and gradually rolling out technology for a national Intranet walled off from the global Internet, called the "Internet E-Paak" or "clean Internet". It routinely throttles bandwidth to slow down connections to virtual private networks and circumvention tools around major events, like elections (Citizen Lab and ASL19 2013; Small Media 2015). Iran has even collaborated with China, and Chinese companies, on its domestic information controls regime. China's ZTE reportedly sold Iranian telecommunications carriers sophisticated equipment capable of monitoring backbone level communications and intercepting emails, and SMS, telephone calls (Stecklow 2012).

Yet Iran also employs a fairly advanced cyber espionage capability that is used to target state adversaries and to gather information on dissidents and human rights campaigners in the global Iranian diaspora. One of the cyber espionage campaigns attributed to the Iranian government, called Newscaster (Ward 2014), exploited several Internet and social media services to target senior US military and diplomatic personnel, congressional personnel, Washington-based journalists, American think tanks, defense contractors in the United States and Israel, as well as vocal supporters of Israel. Newscaster worked by creating fake social media accounts, linking to targets, and then sending spear-phishing emails containing documents embedded with malicious software, which were then used to harvest private email and log-in credentials. Citizen Lab researchers have documented a similar Iranian-based spear-phishing campaign to trick users into giving up their credentials to Gmail accounts, even bypassing Google's two-factor authentication security measures (Scott-Railton and Kleemola 2015).

The actions of Iran in cyberspace are thus a continuation of what Iran has long been doing in more conventional terms. For example, as part of clandestine intelligence support for the Assad regime in Syria, Iran has likely assisted in the organization of targeted digital attacks on the opposition (Regalado, Villeneuve and Scott-Railton 2015). Alongside the flow of finances, weapons, and strategic intelligence to Hezbollah, Iranian intelligence may also have supplied eavesdropping and other information warfare technology leading up to and during the 2006 attacks on

Israel (Cordesman, Sullivan and Sullivan 2007; Wege 2012). After American and Israeli-organized Stuxnet targeting of its nuclear centrifuges, Iran may have repurposed the same malware to target the computers of Saudi Arabia's Aramco refineries (Zetter 2014). To depict Iran as a model of Cyber Westphalia thus obscures the extent to which it has its own elaborate outward-facing digital strategy.

Also like China, as much as Iran wants to limit and contain the free inward flow of information, it depends on transborder communications for a myriad of commercial exchanges (Howard, Agarwal and Hussain 2011). Consider the practical trade-offs confronting Iran in its efforts to throttle access to certain VPNs used to circumvent Iranian firewalls. Traditionally, such circumvention has come at a price: connections to banking and other financial services using the same encryption protocols have been disrupted, to the chagrin of Iranian businesses and elites. Researchers have therefore observed Iranian information controls becoming much more fine-grained and precise, targeting the specific protocols associated with popular VPNs while limiting collateral damage to https connections associated with financial exchanges. This evolution of information controls shows both a maturation of techniques but also clear evidence of the importance of both licit and illicit trans-border traffic to the Iranian economy. Actual Iranian practices suggest a nuanced balancing act, but a robust and deepening international engagement in cyberspace.

Russia presents a similar case. Under the reign of Vladimir Putin, the country has gradually reverted to authoritarian rule, part of which includes a tightening grip on information within Russian territory. A major impetus behind these controls was the 2011 anti-government protests, organized through social media, which took Russian authorities by surprise. In order to contain future demonstrations of this sort, Russian authorities pressured Internet companies to comply with Russian government policies. Today, Russia evinces all of the elements of "Cyber Westphalia" – sweeping data localization laws imposed on foreign Internet giants like Facebook, Google, Twitter, and LinkedIn, a broadening Internet censorship regime, arrests and intimidation of independent media and bloggers, and an architecture of wholesale mass surveillance undertaken by the installation of equipment at telecommunications companies, known as the SORM system (Soldatov and Borogan 2015). Russia and China, moreover, have cooperated on information controls: in April 2016, Russia hosted the first Russia-China cyber security forum to share strategies and best practices. The meeting included Lu Wei, head of China's State Internet Information Office and Fang Bixang, the man widely thought to be the "father" of China's Great Firewall.

As in the cases of China and Iran, however, Russia's information controls are not limited by its territorial boundaries. Russia's approach to cyberspace is instead highly elaborated. It is a key part of a larger geopolitical strategy that includes industrial scale cyber espionage and targeted digital attacks, sophisticated propaganda and disinformation campaigns through state-controlled media organs, and the extension of Russian equipment, technology and know-how to former client states, particularly in the countries of the former Soviet Union. For example, many members of the Commonwealth of Independent States have in place a SORM-compliant system of

mass surveillance, the technical equipment for which is shared by Russian security services. Russian manufactured telecommunications routers are deployed throughout Asia and may contain hidden surveillance functions engineered by design to allow Russian interception. CIS countries also coordinate their cyber security strategies through regional forums like the Shanghai Cooperation Organization, the SCO, which also includes China, Iran, and Pakistan. The SCO have developed collective approaches to repelling social media inspired protests, which are typically framed by the rubric "counter-terrorism".

Russia is widely considered to be a tier-one cyber espionage power connected to many international cyber espionage campaigns. It is assumed, moreover, that Russian SIGINT makes use of the talented organized criminal groups that have long flourished in Russia and whose skills are connected to thriving science, technology, engineering and mathematics programs. The use of organized crime for offensive cyber operations is a convenient way to reap the benefits of such attacks while providing a cloak of plausible deniability, as evidenced in Estonia in 2007 and Georgia in 2008 (Deibert, Rohozinski and Crete-Nishihata 2012).

What we do know about Russian SIGINT campaigns is indicative of extraordinary skill at leveraging a multitude of mostly free Internet services to reach far across global cyberspace to gather information. Consider the so-called "Turla Group" Russian cyber espionage campaign, which affected many high-value targets in dozens of countries worldwide, and which uses earth-orbiting satellite uplinks as command and control servers (Tanase 2015). Another sophisticated Russian cyber espionage campaign (FireEye 2015), referred to in the security industry as APT29, uses a digital mélange (Lennon 2015) of Internet tools, like Twitter, Github, as well as file sharing and cloud computing services, to distribute its command-and-control infrastructure and help obfuscate the identity of those ultimately responsible. Military incursions into Crimea and Ukraine in 2014 and 2015 more directly illustrated an ability to maneuver through cyberspace at will, monitor activities, and mount targeted but isolated malware attacks meant to confuse, weaken, and compromise Ukrainian adversaries. At the same time, Russia had little incentive to disrupt Ukrainian telecommunications systems entirely. Even the December 2015 attack on Ukraine's power grid, which caused a massive power outage and which was attributed to Russian-based hackers, was limited in scope and scale.

One of the distinct traits of Russian cyber espionage are its "influence operations", which have a long history connecting back to the Soviet period. They are digital variations of Cold War propaganda, disinformation, and other espionage campaigns. For example, Russia makes extensive use of social media to discredit and sow discord among adversaries, including the use of paid "trolls" who post messages favourable to the Putin regime, or harass those who are in opposition. Its long-standing use of "Kompromat" – "compromising material" - is commonly used as a technique to discredit political opponents with embarrassing information typically acquired clandestinely and then published. This was taken to a new level with intrusions into the email networks of prominent Democratic Party officials in 2016. Information acquired by Russian-backed cyber-criminal organizations

was then provided to WikiLeaks and other social media, ostensibly to discredit Presidential-candidate Hillary Clinton. While Russia promotes territorially-based information controls in the international sphere, and routinely censors and monitors the Internet and social media within Russian territory, these and other well-publicized influence operations demonstrate that it also actively engages social media and other digital assets abroad in pursuit of its strategic objectives.

Perhaps the countries one would expect to be the least likely to project power extraterritorially would be lower-tier authoritarian, mixed or hybrid regimes and countries, like those in the Gulf, sub-Saharan Africa, the Middle East and North Africa, Asia, Latin America, and the former Soviet Union. Countries like Saudi Arabia, UAE, Bahrain, Sudan, Ethiopia, Egypt, Syria, Vietnam, Thailand, Singapore, Pakistan, Myanmar, Venezuela, Uzbekistan, Tajikistan, Kyrgyzstan, and Kazakhstan might arguably be expected to be principal proponents of a coming Cyber Westphalia. All of these countries have in fact moved aggressively to control domestic information space through technical and regulatory means, and in every case have in place Internet censorship systems to block access to information that crosses their territorial borders. They are also considered principal supporters of Russian and Chinese-backed international initiatives on "Internet sovereignty", and many of them have introduced data localization regulations. But the actual governance practices of a widening range of autocratic countries of the global South are not confined inside their territorial boundaries. Diaspora communities living in the industrialized north use telecommunications networks to send billions of dollars of remittances to back to their originating countries. These diaspora networks also organize politically in ways that may challenge autocratic regimes. Holes in digital firewalls are exploited to advance human rights campaigns and support independent media outlets. At the same time, autocrats themselves continue to depend on open extraterritorial communication channels to bolster their rule or prepare for the future by way of offshore banking havens and real estate investments. While their strategic aspirations may not match those of the United States, such activities require the external projection of digital power.

Finally, authoritarian countries constitute a growing and profitable client base for a vast and rapidly expanding cyber-security industry, which can help control information within territorial boundaries and assist in efforts to investigate and neutralize threats abroad. Thanks to the commercial spyware industry, for example, some of the world's least connected and most impoverished countries, which lack domestic science, technology, and mathematics capacities, are nonetheless able to purchase their own sophisticated "NSA"-like capabilities. Off-the-shelf digital tools are readily available from companies headquartered in the west. Citizen Lab researchers have mapped the proliferation of such commercial spyware services to dozens of authoritarian regimes in all regions of the global South. Espionage operations undertaken using these services typically target diaspora networks. They can be routed through multiple state jurisdictions to obfuscate their origins. Even US-based cloud-computing infrastructure is now routinely employed in the espionage operations

of intelligence agencies based in the global South (Marczak et al. 2014; Marczak, Scott-Railton et al. 2015; Marczak, Scott-Railton and McKune 2015).

Conclusion

State agencies around the world are energetically attempting to re-establish territorial control over the Internet. At the same time, they are increasingly engaged directly or indirectly in extraterritorial projections of power in and through cyberspace. Their mutual entanglement both expands and constrains their own strategic options. The effective sovereignty of states defined in terms of uncontested territorial control in this domain has always been an illusion. The contested openness of cyberspace today, however, exposes the extent to which offensive and defensive policies have in fact constituted a new and very highly interdependent systemic architecture. Anarchy does not describe its political underpinnings, and neither does a straightforward notion of hierarchy. The interaction of dominant intelligence-sharing arrangements of rapidly expanding scope, challenges from autocratic governments simultaneously threatened and empowered by digital openness, and the rapid technological deepening of transnational networks permissive of the nearly instantaneous transmission of data –all render authoritative rule by states ever more complex in principle and in practice.

American hegemony in cyberspace was once manifested by governmental agencies and American-led firms and non-state actors. The observable fact today is that even core interests of the United States cannot be secured without the active collaboration of a growing community of allies and a degree of acquiescence and self-restraint by challengers. The currently surging extraterritorial exercise of both official and corporate digital power, moreover, entangles the United States and other public authorities around the world even as it transforms and reshapes cyberspace itself.

For every "Internet blackout" or "national Intranet" researchers identify, they also find the regimes behind them exploiting transnational communications systems and using common protocols to infiltrate adversaries, gather intelligence, and influence and shape events outside their territories. Together, these kinds of activities have a combined network effect, continuously re-embedding political authorities in distributed and fast-changing digital webs. Extraterritorial projections of state power contribute to a mutual entanglement that has collectively channelled and localized conflict, while restraining temptations to engage in all-out electronic warfare (Lindsay 2017). States continue to depend on and benefit from global networks, and even the most autocratic of them confront compelling incentives not to disable or destroy them (Mueller 2010). Despite all the attention recently paid to issues of digital attacks, disinformation campaigns, and money laundering, researchers have uncovered no suggestion of an emerging global consensus that would be required to secure impenetrable boundaries in cyberspace. The absence of such evidence, moreover, does not appear simply to be attributable to the momentary material interests of dominant social and political elites.

The openness of cyberspace is endogenous to all dimensions of national security policy, the traditional core feature of territorial states.

Watch what actual state agencies do, not what they say. A paradox is revealed. Persistent competitive pressures incentivize attempts to manipulate telecommunications networks internally and externally, but taking those global networks down is widely understood to be self-destructive and self-defeating. Territorially-anchored states depend on trans-border networks to defend themselves and to project their power abroad. The more states become entangled in those networks, the less likely they are to degrade or destroy them, and the more likely they are to join overtly or tacitly in common cause if a rogue non-state actor seriously threatens them.

The broader implications of mutual entanglement in cyberspace bear on the changing character of global political authority itself. National, intergovernmental, and transnational forces together determine the contours of the very space within and through which states now act. Even dominant states must live with the structural denial of locality in this critical domain. The essential quality of cyberspace binds them. Indeed, the kinds of evidence outlined in this chapter suggests that cyberspace is having a transformative impact on the territorial state as conventionally conceived. In the end, global networks cannot be effectively and legitimately governed at the national level. A process of unbundling political authority and recasting it is underway. That process reflects the dynamic interplay of frustrated impulses toward re-territorialization and the imperatives of projecting power extraterritorially. A high degree of global policy ambiguity may therefore be expected for the foreseeable future, since multiple and overlapping claims over rights and responsibilities in cyberspace look set to remain in contention. The external projection of power in and through cyberspace, nevertheless, disturbs systemic order and forces observers to contemplate the entangling effects of functional and political spill-overs (Braman 2013; Daskal 2015; Mueller 2017).

Mark Zacher long ago underlined the transformation underway within and among states in a system that inclined toward openness and encouraged deeper and more intrusive "violations" of Westphalian sovereignty: "cobwebs of agreement have grown, and states have become more aware of the importance of both order and openness for national prosperity" (Zacher and Sutton 1996, 232–3). Of course, the leaders of states have always adhered formally to conventional norms of sovereignty, but they have also often deviated when necessary or convenient (Krasner 1999). Mutual entanglement in cyberspace suggests both mounting constraints on such tactical political calculations as well as stark disappointment for libertarian hopes of Internet freedom. The transformative complexities of sovereign authority when a still territorially anchored political system meets an increasingly global social and economic system are observable in cyberspace (Buzan and Lawson 2015; Grande and Pauly, 2005; Rosenau 2003; Zürn 2018).

At the global level, it is not difficult to discern the outlines of emergent "digital security communities". More implicit but also becoming discernible are workable understandings among strategic competitors. Despite very different perspectives

on the meaning or applicability of the rule of law and the supposed sanctity of the principle of non-interference in internal affairs, China and the United States, for example, have begun discussions on acceptable behaviour in cyberspace. They are doing so bilaterally as well as multilaterally through the United Nations and other forums (G20 Research Group 2015; United Nations 2015). The two countries even agreed in September 2015 on a limited set of restraining principles concerning industrial espionage, the theft of intellectual property, the targeting of critical infrastructure, and the need to cooperate on investigations of electronic crimes (Harold, Libicki and Cevallos 2016; The White House 2015). Despite continuing concerns about compliance, other countries are moving in the same direction as they contemplate expanding Chinese investment across a range of economic sectors. Future on-line activities, and the cooperation of like-minded authorities on either side of the autocracy divide, will determine whether new principles and policies become effective and broadly accepted as legitimate.

The risks of catastrophic miscalculations or accidents in cyberspace remain. But uncertainty at this point does not suggest disorderly fragmentation along territorial lines. Through their trans-territorial interaction, the governments of states will likely continue to live with paradox. They look set to continue moving toward new forms of authority to govern the Internet – contested but acceptable enough to permit a dynamic system to persist. States still matter, but mutual entanglement in cyberspace reinforces the idea of complex sovereignty and mocks dreams of retreat to a simpler past.

Note

1 Portions of this chapter are derived from Ronald J. Deibert and Louis W. Pauly, "Cyber Westphalia and Beyond: Extraterritoriality and Mutual Entanglement in Cyberspace", paper prepared for the Annual Meeting of the International Studies Association, Baltimore, Maryland, February 2017, and Ronald J. Deibert and Louis W. Pauly, "Boundaries and Borders in Global Cyberspace", in *Borders, Boundaries, and the Future of Canadian Society*, The Third Annual SD Clark Symposium on the Future of Canadian Society Celebrating Canada's Sesquicentennial, 10 November 2017. For comments, we thank Daniel Deudney, Joseph Nye, Abe Newman, Jon Lindsay, Lennart Maschmeyer, Hans Klein, Milton Mueller, and the editors of this volume. We also gratefully acknowledge support provided by the Canada Research Chairs Program and the Social Sciences and Humanities Research Council of Canada.

References

Blue, Violet. 2012. "Hack in The Box." ZDNet, October 11; accessed at: http://www.zdnet.com/article/hack-in-the-box-researcher-reveals-ease-of-huawei-router-access/, May 13 2016.

Braman, Sandra. 2013. "The geopolitical vs. the network political." *International Journal of Media and Cultural Politics* 9(3): 277–296.

Buzan, Barry, and George Lawson. (2015). *The Global Transformation*. Cambridge: Cambridge University Press.

Citizen Lab. 2014. "Communities@Risk." Citizen Lab, Report 48; accessed at: https://targetedthreats.net/index.html, May 13, 2016.

Citizen Lab and ASL19. 2013. "After the Green Movement." OpenNet Initiative Special Report, March 11; accessed at: https://opennet.net/sites/opennet.net/files/iranreport.pdf, May 13 2016.

Cordesman, Anthony H., George Sullivan, and William D. Sullivan. 2007. "Lessons of the 2006 Israeli-Lebanon War." Center for Strategic and International Studies, Significant Issues Series, 29(4); accessed at: https://csis.org/files/publication/120720_Cordesman_LessonsIsraeliHezbollah.pdf, May 13 2016.

Dalek, Jakub, et al. 2015. "A Chatty Squirrel: Privacy and Security Issues with UC Browser." Citizen Lab, Research Brief 55, May 21; accessed at: https://citizenlab.org/2015/05/a-chatty-squirrel-privacy-and-security-issues-with-uc-browser/, May 17 2016.

Daskal, Jennifer. 2015. "The Un-Territoriality of Data." *Yale Law Journal*, 125(2): 326–398.

Deibert, Ronald. 2012. "Social Media, Inc.: The Global Politics of Big Data." *World Politics Review*, June 19; accessed at: http://www.worldpoliticsreview.com/articles/12065/social-media-inc-the-global-politics-of-big-data, May 17 2016.

Deibert, Ronald. 2015. "Authoritarianism Goes Global." *Journal of Democracy* 26(3): 64–78.

Deibert, Ronald. 2017. "Cyber Security." In *Routledge Handbook of Security Studies*, edited by Myriam Dunn Cavelty and Thierry Balzacq, 172–182. Abingdon: Routledge.

Deibert, Ronald, and Masashi Crete-Nishihata. 2012. "Global Governance and the Spread of Cyberspace Controls." *Global Governance* 18(3): 339–361.

Deibert, Ronald, et al., eds. 2008. *Access Denied: The Practice and Policy of Internet Filtering*. Cambridge, MA: MIT Press.

Deibert, Ronald, et al., eds. 2010. *Access Controlled: Policies and Practices of Internet Filtering and Surveillance*. Cambridge, MA: MIT Press.

Deibert, Ronald, et al., eds. 2012. *Access Contested: Security, Resistance, and Identity in Asian Cyberspace*, Cambridge, MA: MIT Press.

Deibert, Ronald, and Rafal Rohozinski. 2008. "Good for Liberty, Bad for Security? Internet Securitization and Global Civil Society," in *Access Denied*, edited by Ronald Deibert, et al., 123–165. Cambridge, MA: MIT Press.

Deibert, Ronald and Rafal Rohozinski. 2010. "Risking Security: The Policies and Paradoxes of Cyberspace Security." *International Political Sociology* 4(1): 15–32.

Deibert, Ronald, Rafal Rohozinski, and Masashi Crete-Nishihata. 2012. "Cyclones in Cyberspace." *Security Dialogue* 43(1): 3–24.

Demchak, Chris C., and Peter Dombrowski. 2011."Rise of a Cybered Westphalian Age." *Strategic Studies Quarterly* 5(1): 32–61.

Deudney, Dan. 2007. *Bounding Power*. Princeton, NJ: Princeton University Press.

Dombrowski, Peter. 2016. "China wants to Draw Borders in Cyberspace." New *Perspectives Quarterly* 33(2): 38–42.

Drake, William J., Vinton G. Cerf, and Wolfgang Kleinwächter. 2016. "Internet Fragmentation: An Overview." World Economic Forum, Future of the Internet Initiative Whitepaper; accessed at: www3.weforum.org/docs/WEF_FII_Internet_Fragmentation_An_Overview_2016.pdf, May 12 2016.

FireEye. 2015. "HAMMERTOSS." FireEye Special Report; accessed at: https://www2.fireeye.com/rs/848-DID-242/images/rpt-apt29-hammertoss.pdf, May 13 2016.

Friends Committee on National Legislation. 2015. "Understanding Drones." Friends Committee on National Legislation, accessed at: http://fcnl.org/issues/foreign_policy/understanding_drones/, May 13 2016.

G20 Research Group. 2015. G20 Leaders' Communiqué', G20 Research Group, November 16; accessed at: http://www.g20.utoronto.ca/2015/151116-communique.html, May 13 2016.

Harold, Scott Warren, Martin C. Libicki, and Astrid Stuth Cevallos. 2016. "Getting to Yes with China in Cyberspace." RAND Corporation, Research Report 1335; accessed at: www.

rand.org/content/dam/rand/pubs/research_reports/RR1300/RR1335/RAND_RR1335.pdf, May 13 2016.

Grande, Edgar, and Louis W. Pauly eds. 2005. *Complex Sovereignty*. Toronto: University of Toronto Press.

Gunitsky, Seva. 2015. "Corrupting the Cyber-Commons: Social Media as a Tool of Autocratic Resilience." *Perspectives on Politics* 13(1): 42–54.

Harris, Shane. 2014. *@War: The Rise of the Military–Internet Complex*. Boston, MA: Eamon Dolan/Houghton Mifflin Harcourt.

Howard, Philip N., Sheetal D. Agarwal, and Muzammil M. Hussain. 2011. "The Dictators' Digital Dilemma." *Issues in Technology Innovation* 13: 1–11.

Johnson, David R., and David G. Post. 1996. "Law and Borders – The Rise of Law in Cyberspace." *Stanford Law Review*, 48(5):1367–1402.

Katzenstein, Peter J., ed. 2012. *Anglo-America and Its Discontents*. Abingdon: Routledge.

Knockel, Jeffrey, Sarah McKune, and Adam Senft. 2016. "Baidu's and Don'ts." Citizen Lab, Research Brief 72, February 23; accessed at: https://citizenlab.org/2016/02/privacy-security-issues-baidu-browser/, May 17, 2016.

Knockel, Jeffrey, Adam Senft, and Ronald Deibert. 2016. "Wup! There it is." *Citizen Lab*, Research Brief 75, March 28, accessed at: https://citizenlab.org/2016/03/privacy-security-issues-qq-browser/, May 17 2016.

Krasner, Stephen. 1999. *Sovereignty*. Princeton, NJ: Princeton University Press.

Lennon, Mike. 2015. "Russian Hacker Tool Uses Legitimate Web Services to Hide Attacks: FireEye." *SecurityWeek*, July 29; accessed at: http://www.securityweek.com/russian-hacker-tool-uses-legitimate-web-services-hide-attacks-fireeye, May 13 2016.

Lindsay, Jon R. 2015. "The Impact of China on Cybersecurity: Fiction and Friction." *International Security* 39(3): 7–47.

Lindsay, Jon R. 2017. "Restrained by Design: The Political Economy of Cybersecurity." *Digital Policy, Regulation and Governance*, 19(6): 493–514.

Mandiant. (n.d.). APT1: Exposing One of China's Cyber Espionage Units, accessed at: www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf, January 2, 2019.

Marczak, Bill, et al. 2014. "Mapping Hacking Team's Untraceable Spyware." *Citizen Lab*, Research Brief 33, February 17; accessed at: https://citizenlab.org/2014/02/mapping-hacking-teams-untraceable-spyware/, May 13 2016.

Marczak, Bill, John Scott-Railton, and Sarah McKune. 2015. "Hacking Team Reloaded?" *Citizen Lab*, Research Brief 50, March 9; accessed at: https://citizenlab.org/2015/03/hacking-team-reloaded-us-based-ethiopian-journalists-targeted-spyware/, May 13 2016.

Marczak, Bill, et al. 2015. "Pay No Attention to the Server Behind the Proxy." *Citizen Lab*, Research Brief 65, October 15; accessed at: https://citizenlab.org/2015/10/mapping-finfishers-continuing-proliferation/, May 13 2016.

Marczak, Bill, Nicholas Weaver, et al. 2015. "An Analysis of China's Great Cannon." 5th USENIX Workshop on Free and Open Communications on the Internet, Washington, D.C.

Marquis-Boire, Morgan, Glenn Greenwald, and Micah Lee. 2015. "XKEYSCORE." *The Intercept*, July 1; accessed at: https://theintercept.com/2015/07/01/nsas-google-worlds-private-communications/, May 13, 2016.

Mozur, Paul. 2015. "Baidu and CloudFlare Boost Users Over China's Great Firewall." *The New York Times*, September 13; accessed at: http://www.nytimes.com/2015/09/14/business/partnership-boosts-users-over-chinas-great-firewall.html?_r=0, May 13 2016.

Müller-Maguhn, et al. 2014. "Treasure Map." *Der Spiegel*, September 14; accessed at: http://www.spiegel.de/international/world/snowden-documents-indicate-nsa-has-breached-deutsche-telekom-a-991503.html, May 13 2016.

Mueller, Milton L. 2010. *Networks and States*. Cambridge, MA: MIT Press.

Mueller, Milton L. 2017. *Will the Internet Fragment?* Cambridge: Polity.

Nye, Joseph. 2017. "Deterrence and Dissuasion in Cyberspace." *International Security*, Winter, 41(3): 44–71.

Ruan, Lotus, et al. 2016. "One App, Two Systems." Citizen Lab, November 30, 2016; accessed at: https://citizenlab.ca/2016/11/wechat-china-censorship-one-app-two-systems/, February 3, 2018.

Ruggie, John G. 1993. "Territoriality and Beyond." *International Organization*, 47(1): 139–174.

Regalado, Daniel, Nart Villeneuve, and John Scott-Railton. 2015. "Behind the Syrian Conflict's Digital Front Lines." *FireEye Special Report*, February 2015; accessed at: https://www.fireeye.com/content/dam/fireeye-www/global/en/current-threats/pdfs/rpt-behind-the-syria-conflict.pdf, May 13 2016.

Rosenau, James. 2003. *Distant Proximities*. Princeton, NJ: Princeton University Press.

Scott-Railton, John, and Katie Kleemola. 2015. "London Calling: Two Factor Authentication Phishing from Iran." *Citizen Lab*, Research Brief 61, August 27; accessed at: https://citizenlab.org/2015/08/iran_two_factor_phishing/, May 17 2016.

Small Media. 2015. "Iranian Internet Infrastructure and Policy Report." *Small Media*; accessed at: https://smallmedia.org.uk/sites/default/files/u8/IIIP_Feb15.pdf, May 13 2016.

Soldatov, Andrei, and Irina Borogan. 2015. *The Red Web*. New York: Public Affairs.

Starosielski, Nicole. 2015. *The Undersea Network*. Durham, NC: Duke University Press.

Stecklow, Steve. 2012. "Chinese firm helps Iran spy on citizens." Reuters, March 22; accessed at: http://www.reuters.com/article/us-iran-telecoms-idUSBRE82L0B820120322, May 13 2016.

Stiennon, Richard. 2016. "The Entire IT Security Landscape." 25th RSA Conference San Francisco, CA, March 18; accessed at: https://www.youtube.com/watch?v=YYNM2VRmncE, May 13 2016.

Tanase, Stefan. 2015. "Satellite Turla." *Securelist*, September 9; accessed at: https://securelist.com/blog/research/72081/satellite-turla-apt-command-and-control-in-the-sky/, May 13 2016.

Union of Concerned Scientists. 2016. "Union of Concerned Scientists Satellite Database." February 25; accessed at: http://www.ucsusa.org/nuclear-weapons/space-weapons/satellite-database#.VyykWxUrKgQ, May 13 2016.

United Nations. 2015. "Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security." United Nations General Assembly A/70/174, July 22; accessed at: http://www.un.org/ga/search/view_doc.asp?symbol=A/70/174, May 13 2016.

VanderKlippe, Nathan. 2016. "China collecting sensitive personal data through Android apps." *The Globe and Mail*, February 24; accessed at: http://www.theglobeandmail.com/technology/tech-news/millions-of-android-apps-send-sensitive-data-to-china-u-of-t-report/article28865055/, May 13 2016.

Villeneuve, Nart. 2008. "Breaching Trust." *Information Warfare Monitor*; October 1; accessed at: https://www.scribd.com/doc/13712715/Breaching-Trust-An-analysis-of-surveillance-and-security-practices-on-China-s-TOM-Skype-platform, May 17 2016.

Vine, David. 2015. *Base Nation*. New York: Henry Holt.

Vodafone. 2014. "Vodafone Law Enforcement Disclosure Report 2013/2014." Accessed at: https://www.vodafone.com/content/sustainabilityreport/2014/index/operating_responsibly/privacy_and_security/law_enforcement.html, May 12 2016.

Ward, Stephen. 2014. "An Iranian Threat Inside Social Media." iSIGHT Partners, May 28; accessed at: https://www.isightpartners.com/2014/05/newscaster-iranian-threat-inside-social-media/, May 13 2016.

Wege, Carl. 2012. "Hizballah's Counterintelligence Apparatus." *International Journal of Intelligence and Counter-Intelligence*, 25(4): 771–785.

White House. 2015. "President Xi Jinping's Visit to the United States." White House Press Release, September 25; accessed at: https://www.whitehouse.gov/the-press-office/2015/09/25/fact-sheet-president-xi-jinpings-state-visit-united-states, May 13 2016.

Zacher, Mark W., and Brent A. Sutton. 1996. *Governing Global Networks*. Cambridge: Cambridge University Press.

Zetter, Kim. 2014. *Countdown to Zero Day*. New York: Crown Publishers.

Zürn, Michael. 2018. *A Theory of Global Governance*. Oxford: Oxford University Press.

6

DIGITAL DATA AND THE TRANSNATIONAL INTELLIGENCE SPACE

Didier Bigo and Laurent Bonelli

Introduction

The Edward Snowden disclosures on the American National Security Agency's (NSA) large-scale digital capture practices have spawned the opening of a series of political, juridical, philosophical, and academic debates. Discussions have predominately counterpoised the relationship between mobility and communications control, on the one hand, and the exponential growth in the amount of traces left by the daily activities of individuals using digital technologies, on the other. But what exactly are traces, what do they record, and how are they being recorded? Are they "raw data" available to all or, instead, data that belong to the realm of the private?[1] To whom do the data belong? To what extent do they constitute new sources of enrichment, awareness, commercial profits, statistical knowledge on populations, knowledge on the intimate lives of individuals, and, of course, surveillance?

If the internet was at one time perceived as the place par excellence for knowledge exploration and the organization of remote encounters, it is now increasingly being seen as a world that exacerbates the expansion of neoliberal capitalist logics. Within the directives of the latter, digital data become a new raw material that is both free and can be used to monitor the activities and behaviours of individuals with the help of automated data collection technologies. Challenging the relevance of national borders, the internet has also been understood as a key vector of globalized communication, wherein anonymity has allowed for networks to be created according to the affinities and mutual interests of individuals. In destabilizing notions of internal and external geography and thereby blurring or superimposing borders, the internet has had an impact on uses of violence, security mechanisms, and intelligence logics.

In exchange for the "free" use of internet services and resources, commercial actors making the internet work and further developing digital technologies believe to have the inherent right to exploit data produced by individuals. As a consequence,

internet users have been subject to a "digital encomienda."[2] Organizations interested in intelligence, in the broad sense of the term—be it the police and the military, or immigration and customs officers—have come to see the internet as somewhat of a double-edged sword. On the one hand, it is a major risk as it devalues their pre-existing professional routines. On the other hand, it presents unique opportunities to gain in-depth knowledge on individual practices, which had previously only been held in the hands of the private sector. These actors, however, have by no means reacted uniformly—either through functional adaptation or through the coordination of an "intelligence community"—to these pitfalls and promises. What they have all done, though, is subsequently integrate the collection of personal data and the analysis of these digital traces into their repertoire of activities.

In some cases, the interception of large amounts of data, with the help of algorithms, allows for the detection of behavioural abnormalities. This information can subsequently be used to identify risk profiles. However, as we will argue in this chapter, the way that these practices are performed vary greatly depending on one's degree of seniority in their occupational field, their capacity in terms of personnel, their technical skills (hardware and software), as well as their subjective visions on what exactly counts as "intelligence." Depending on their practical goals and know-how, intelligence agents produce different interpretations of what these influxes of data and analytical treatments can do for their profession.

These various actors first discussed amongst themselves and with politicians the value that data constituted by traces left on the internet, potentially amassed, and linked together using database software in order to generate statistical information might have for intelligence activities. This then raised the question of the utility of dedicating significant financial and human resources to the acquisition of remote interception technologies (satellite, digital) and their relative advantage in comparison to human means that could be used to reach the same results, notably by employing undercover techniques and informants. As we shall see, individual and collective actors responded to this dilemma quite differently. If various actors situated in the "field of intelligence professionals" have come to realize how easy it is to accumulate, exchange, and store digital data, to what extent has this accumulation of data been counterproductive, leading these professionals to miss the specific forest for the millions of trees?[3] The most specialized services (intelligence-counterintelligence) are still not convinced by accumulation techniques (collect it all) and have instead preferred to keep their sensitive case files outside of shared collection and exchange circuits. The outbreak of a number of public controversies—notably following the disclosures on the human rights violations committed by the Central Intelligence Agency (CIA) and its accomplices, the large-scale data capture practices of the NSA and the "Five Eyes," (United States of America, United Kingdom, Canada, Australia, New Zealand) and the near-routine use of drones outside of active conflict zones—forced intelligence agencies to rethink the utility and value of these strategies and devices. Digital technologies were questioned not only in terms of their capacity to provide greater security, but also in terms of the legal questions their use raised, privacy

issues, and, more generally, their adequacy with the position-taking in terms of values of countries that claim to be democratic and contest the practices of authoritarian regimes (Bigo 2012, 2016).

In this chapter, we will examine only some of the agents situated in the intelligence field of sensitive information. On the one hand, these professionals belong to intelligence agencies that see themselves as working to defend the national interests of their country. On the other hand, these professionals also exchange data with their counterparts in the national security agencies of other countries. These intelligence service agents thus have the capacity and the authority to intercept data not only at home but also abroad. For the most part, represented countries include former colonial and neo-colonial powers of the Global North, who esteem that they have a role to play at the regional or global level. Together, these agents form a transnational space that is linked by virtue of historical alliances that were first established during World War II and, more recently, through the efforts of those who have come to play a major role in the geopolitics of Internet cables.[4] This space has been named after a group specializing in communications surveillance: the Five Eyes, or the Five Eyes Plus. However, as we shall see, this transnational intelligence space is not limited to the intelligence agencies that are members of that exclusive group. The story of the alliance that lead to the creation of the Five Eyes is quite well known but has often been summarized as a story about common sensibilities shared between intelligence agencies with Anglo-Saxon origins, that created the necessary conditions for a form of mutual trust to develop between political leaders and agency actors. However, we are not convinced of this historical-cultural narrative on trust established between similar countries. Such an argument implicitly assumes that each country has a clear national history as well as a homogenous intelligence policy, meaning that the only thing that the researcher would need to do, is to compare these national trajectories to understand how trust first emerged between the concerned countries. Instead, we propose a study of the means and practices of intelligence agencies in order to then chart their position within a transnational space, without assuming that national or cultural belonging creates positions of proximity between agencies. In determining our case selection based on power relations, our focus will be on intelligence agencies that are the most resource-endowed in quantitative terms, that demonstrate a degree of professionalism, and that have a long history of managing sensitive information. Inspired by the work of Pierre Bourdieu, our study seeks to systematize elements collected in interviews with intelligence professionals by employing a structural analysis of the space in which the selected intelligence agencies are situated. To do so, we perform a multiple correspondence analysis (MCA), which allows us to rigorously visualize the space of institutional positions based on a series of defining characteristics (type of missions conducted, supervisory authority, territory of action, staff numbers, technological capital, etc.). In making connection between these objective positions and the discourses of actors regarding their practices and the meaning of intelligence, we are able to identify homologies as well as divergences that structure cooperation and data exchanges between agencies.

As we shall see in this study, it is not the number of internet traces left behind that matter. The fact that traces are produced does not automatically turn them into a source of wealth or power. Instead, what matters is how such traces are constituted as data and used for intelligence policy purposes as well as the horizon of suspicion in which they are used. So why intercept data? Simply because they are available and can be "picked" as flowers growing in a free space? Should we do this for all data in order to have a comprehensive graph that represents relationships between individuals and large groups of the population? Or should we restrict our use of data and leave them where they are, thereby preventing their use in cross-checking?

The existence of an intelligence policy that involved the large-scale surveillance of masses of individuals, categorized as suspect or as undesirable, has for the most part been trivialized. It has been argued that linking intelligence techniques with automated technologies that record digital traces left by the activities of individuals and their transactions is justified when it's preventive and protective function can help to anticipate and avoid violence. Yet, in our opinion, the relationship between the existence of the digital and predictive intelligence is not in and of itself inescapable. As opposed to being due to the inherent nature of the technology, this association has been politically modelled in a specific international context and depends on power struggles between the actors who determine the use and exchange value of digital data. The value of data is determined by the degree to which they can generate suspicion—notably when it comes to future acts—even if correlations made are so weak that they do not hold when faced with the law. Thus, markedly in contrast with legal practices, the actors of this space of doubt, of suspicion, and of possibilities play the role of "prince counsellors" that provides advice before decisions are made by politicians. Moreover, the symbolic value of intelligence data depends less on its content—despite the ideology of secrecy that sanctifies this content—than it does on who produced it, in what context, and for what reason.

It is this last point that we will deal with more substantially as it was paradoxically concealed in general statements made on the surveillance "society" and on algorithmic reasoning, which incorrectly suggest that internet users from around the world have been complicit in their own voluntary servitude (Bauman and Lyon 2013, Lehr 2019). This requires us to think reflexively about what the term "intelligence services" (or "security services") really means and the relations between the practices of intelligence agencies on one side and the modalities of digital data surveillance on the other side. Though fairly common in international relations, this chapter will not provide a disembodied analysis of intelligence practices or a history without actors, where intelligence agencies are seen as obeying the orders of political leaders who determine overarching strategies. Instead, we will highlight the heterogeneous characteristics that define intelligence actors, pointing to their differences in terms of socialization, professional habitus, and of different types of missions and actions that are performed. In doing so, we will identify arcs of tension that exist between organizations whose logics of action and modes of reasoning are either antagonistic or, at the very least, advance opposed strategies.

Our analysis will thus shed light on power relations that, thus far, discussions on rivalries between services have failed to capture.

Data, information, intelligence: data as performances and products of competition between intelligence agencies

What do we call "data" when this terminology is used for and in relation to political intelligence purposes? How are data generated and integrated into information chains that allow for the production of analyses that respond to the demands of politicians? What place then does this kind of data occupy in what scholars have termed the "intelligence cycle"? (Gill and Phythian 2016, McElreath, Graves, and Jensen III 2017, Murphy 2016).

Can the data be described, as some believe, as constituting all the traces of a person's or group's activities that may have been collected automatically or intentionally and which are then grouped into files? Referred to as "raw" data, does the data contain generic information in terms of the location of a person associated with an event, at a given moment in the past or in the present? In a second step, can this information then be used to anticipate future behavior through the application of algorithmic software? Within the data, can a distinction be made between content data, which reveal personal opinions based on content and so-called "connection" data, that is "metadata" or tracking data that make associations between individuals based on the exchanging of messages or the sharing of websites as well as through the establishment of shared interests based on the frequentation of the same people and places? This distinction between content and connection data has been presented by many agencies as a technically-relevant difference, as there would be only limited constraints in the exploitation of the latter in comparison to the former.[5] This distinction appears in a number of reports and analytical documents, notably in the United States. However, in opposition to this impersonal interpretation of technical data, various European Courts have pointed out that all tracking and localization data, whether derived from content or connection data, interfere with the privacy of individuals, and as a result are protected by international laws and agreements on personal data.[6] As we can see in these debates and developments, the issue of data ownership is absolutely crucial, as are the ways in which data are created and used for different purposes. Based on this observation, it would then seem necessary to reverse the dominant thinking about data. In other words, data are not the sources of information and analysis, but they are instead the product of it.

Data ownership: an electronic encomienda

The issue of data and the definition of this term cannot be settled by way of a technical consensus. It is a political and legal controversy, which necessarily undermines any conception of raw data as simply technical property that keeps track of the flow of information and to some extent to origin of this information, but that is independent from the ends for which it is used.

As we will argue here and have done so elsewhere, it seems that, on the contrary, it is exactly the different purposes for which data are used that play a role in the construction of the meaning and the form that data take. These meanings and forms are not natural, nor are they raw. Rather, they are the product of specific performances done by a series of actors. This standpoint, however, is not always recognized or appreciated at its face value. In interviews with actors coming from various intelligence agencies as well as in the narratives of scholars working on the "cycle" of intelligence (i.e. intelligence studies), physiocratic and industrialist visions are often mobilized as metaphors when describing the nature of data. For example, in physiocratic analogies, data is represented as a flower or vegetable that awaits harvesting—it is sown by internet users themselves, randomly moved, and left idle or exchanged for services provided by private companies, thereby no longer belonging to the sowers. The data is collected like the celestial manna, granted not by a divine figure but instead by computer science. In industrialist depictions, data is compared to a precious mineral that can be extracted from veined ore rock. In such an analogy, it becomes important to have the right drilling tools that can detect what is important and consequently select and retain only what is of value. Given the mass amounts of heterogeneous and weakly correlated data that circulate, it would be necessary to capture, intercept, and trace data that correspond to a specific profile. Ideally, information would emerge from connections made by that profile, which could then be refined and cut, like diamonds. The desired output would be analysed and that would lead not only to quality information, but to useful information that can be mobilized in political decision-making processes. The transformation of data into politically-relevant information is performed through the analytical practices of intelligence professionals and defines their very métier, which involves much more than algorithmic statistical correlations or the idea of simply collecting information that is already "out there." The two metaphors therefore are not so much about the way these professionals work but are instead used to suggest that the "raw" data do not belong to anyone and are therefore there for the taking. In both visions, individuals are not seen as having ownership over their data. Instead, data are available to those who exploit them and don't have value in and of themselves but acquire added value for those that make data connections and articulations. Data only "make sense" when information is extracted. Taking stock of this overall process, we argue a digital encomienda is at work.[7] As during the Spanish colonization, the "natives" (here, the internet users) have been deprived of their ownership rights and of their status as citizens of the worldwide web. This creates the conditions of possibility for the "colonization" of the web to generate profits and intelligence data. In exchange, web users receive the benefits of more targeted marketing and consumption, remote contacts and friendships (i.e. Friendship 2.0), and allegedly protection against terrorism.[8]

However, while intelligence agencies may be in favor of this primitive political economy of data, these data have an origin. As European courts and data protection authorities have repeatedly pointed out, these data have initial owners. The drafting and recent implementation of the General Data Protection Regulation (GDPR) in Europe confirms this.[9]

Intelligence data: the work and competitions of intelligence actors

Justifying a series of interception and retention practices, "intelligence studies" theorizations of the "fabrication of information" and its transformation into intelligence as being part of a "cycle" of production essentially aim to naturalize the existence of data along with the right to exploit and aggregate them. This is done in relation to modes of reasoning that are often already constructed, meaning that data are used to confirm these modes of reasoning, not invalidate them. In opposition to this argument, we suggest that "intelligence" data are constructed in a performative manner by the very political decisions that initiate data searches, the social use of surveillance techniques that may or may not render something "visible," and, lastly, the languages used by recipients (multiple, single, unwanted) to encode and decode intercepted data.[10] The performances of intelligence actors are thus dependent and based on data belonging to individuals. Very often, however, these actors colonize individual data and transform them into "intelligence tools" by serializing, anonymizing, and grouping data into files. Data from the intelligence world only becomes data when a political interest in their production and preservation has been established, when decisions have been made on where to draw boundaries, which visible elements should be thrown out, and which traces should not be of interest. Data is produced with the aim of creating lists of threats, risks, and vulnerabilities and identifying suspects. They are creating Data Suspects.

In terms of "data politics," our vision is to insist that data is a special performance that reconfigures the relationship between the digital and the material, while influencing contemporary relations between intelligence, surveillance, violence and obedience. These relations and configurations depend on the internal games of intelligence actors and the way that they define security, insecurity, and fate (Bigo 2008).

From such a perspective, intelligence data is understood as the product of political manoeuvres that constitute these data from the very outset according to that which they are "supposed to see," as though this was a neutral and objective act allowing politicians to make decisions. Yet, the act of creating data by orienting them so that they may prove the (legitimate) suspicion of different connections is very political—not in the decisional but constitutive sense. Moreover, this creative process and its outputs rely on mechanisms of association, connection, filtering, and profiling, which end up categorizing some individuals as more suspect than others, more undesirable than others, and more threatening to the established order than others.

As a result, intelligence data are very rarely sources that permit for the establishment of causalities. Instead, they are the result of a process that seeks to legitimize or delegitimize suspicions held by intelligence actors, which rely solely on correlations and not on evidence, hence structural struggles and oppositions between judicial authorities and intelligence services. Intelligence agents make interpretations and draw portraits that create a form of spectacle, that these agents enact

with their lists of suspects and their operational analyses of possible futures. These interpretations represent the core of the files on which the various intelligence services are working.

One major difference, however, is that while these files were previously materially written and recorded on paper, they are now are digitally written and stored on computers. Does this material shift affect the way files are constituted and the modes of reasoning harbored by intelligence agents? Following the new materialism turn, some authors like Marieke de Goede argue that a profound transformation in the modes of reasoning that are at work in the construction of data, but this is not certain.[11] This new mode of reasoning only seems to touch at the fringes of the intelligence craft, observed amongst services that deal with the mobility of travelers or suspicious financial operations, but not so much amongst intelligence services themselves. As Laurent Bonelli and Francesco Ragazzi (2014) have pointed out, the conjectural reasoning described by Ginzburg (1980) remains fundamental to the practices of numerous agents that prefer the "low tech." So, it is not clear whether an "algorithmic" reasoning—which is based on large-scale correlations and a speculative reasoning specific to computer-based tools, instead of precise chains of causalities—could be opposed to and thought of to fully replace a conjectural reasoning. Drawing from our recent work, the dominant reflection and practice of intelligence agents continues to be organized around files or archives that are shared only by a small group of professionals. These documents are read by groups in connection to other information that is deemed secret and operational. The value of information derived from digital data is far from equal to that derived from "low tech" files. While some digital data may be useful for identification and localization, thereby getting past the quasi-structural anonymous character of the Internet, for many actors, the accumulation of heterogeneous data may contravene the understanding of actions. A reasoning based on algorithmic correlations is not able to capture individual targets, but instead creates culpability based on association. This is why a conjectural-based mode of reasoning may be complemented with speculative information and possibilities, but it will continue to remain the core of a profession that ultimately works to study, discipline, and eventually chastise individuals. Without this practice of indexing, the professional practice of intelligence would comprise of no more than the production of geopolitical generalizations of tendencies, the projection of futures yet with no operational capacity.

Though critical in many respects, a considerable amount of the surveillance studies literature has come to consider too quickly the production and traceability of data as an inevitable feature of modern society. This interpretation, however, tends to approach data as "flat," "rhizomatic," and constituted by the "exhaust data" diffused between all social and international worlds. While traceability may be automatized and facilitate rapid communication, some authors overgeneralize the characteristics of digital data, thereby overlooking the data constitution process, which can make them particular, fragment their meanings, or lead to their integration as political products. These studies thus homogenize the policies and practices

of different worlds of intelligence, ignoring vertical hierarchies, competitions, and differences in the technical methods of surveillance that they use (Amoore 2013, de Goede 2012).

Some monitoring methods are transversal and capable of making the use of such techniques more horizontal by facilitating the remote transmission of information. However, this does not necessarily result in the homogenization of the resources and logics of action driving what the agents are looking for in practical terms. Existing theories of electronic surveillance do not sufficiently distinguish between social worlds, which then leads to the assumption that the digital world has a uniform effect that is determined by the technology itself. For example, the production of digital data in the worlds of health, of international commerce, and of security do not generate the same effects in each respective social world, nor does it determine fixed relations between those worlds. The way that actors use new technologies depends on their past dispositions as well as on their capacity or willingness to transform paper data into computer data. With regard to the latter, this maneuver depends as much on one's technical capacities as it does their views on the importance of secrecy, confidentiality, and interest, which may dissuade some actors from stocking data or disseminating them. Data from the digital world thus affects forms of power and everyday politics, but the opposite is also equally true. Data are integrated into social worlds and into everyday practices only if they are seen by actors as helping them in their power struggles. This is best observed when new technologies are modelled or articulated in relation to existing customs and functions.

In sum, computer technology does not "revolutionize" technology so much as it moulds and adapts itself according to differentiated registers that are brought into the routine practices of different social worlds, which together constitute political intelligence. Moreover, computer technology can in fact reinforce divisions between actors and intensify existing points of tension by favouring certain actors over others. This privilege emanates from a general acceptance of the technical dimension of intelligence, whereby large-scale surveillance will be done remotely, with greater attention given to trend analysis that takes on a preventive and predictive dimension via use of "data derivative" operations.[12] These technological innovations thus favor actors who both use it to revitalize the contributions of agents and operations in the field and to value the accumulation of heterogenous data, which can nevertheless highlight correlations and the human mind would not have been capable of identifying by running algorithmic analyses of big data. This is at least what we have observed in narratives used by the agents who are most interested in defining intelligence as the anticipation of hostile acts, no matter where they come from. This mode of reasoning is probably seen as convincing for new entrants into the field of intelligence, as they can only act at a distance and don't have any field agents on the ground or "relays." However, this understanding of intelligence is far less convincing among an older generation of actors, who have operational capacities and who think in terms of adversaries, enemies, and possibly suspects. This second group of individuals is also quite suspicious of the deindividualization of crime, theologies on the possibility of knowing the

unknowable, and efforts to plan the future as if it were a future perfect. These symbolic struggles over the value and manner intelligence should be done then then determine what data "is."

In the next section, we will further study three professional intelligence groups by analyzing the structural modalities that define them as distinctive spaces, each defined by differentiated modes of socialization. Yet, each of these spaces is subject to the reformulation of their practices due to the socio-technical stakes of the digital, while simultaneously maintaining the ability the frame and structure the way that intelligence data is defined and practiced. So, when read in relational terms, some groups of actors are more or less distinct. When their work practices are transformed, actors are pushed to rethink their identities and power positions. In some cases, these reconfigurations may even challenge the strong felt sense of some that they live in a small world apart from the rest (i.e. the deep state), that has its own rules; a world wherein the use of violence in the name of state reason, secrecy, and impunity vis-à-vis the rule of law are the norm, and wherein these actors are committed to a sense of responsibility and loyalty to specific values that must be safeguarded and yet constantly adapted to practical challenges.

The transnational space of intelligence: the structural modalities and dispositions of actors regarding the digital

Intelligence studies has to a large extent suffered from a form of methodological nationalism that presupposes the existence of a national intelligence community, that collectively defends national interests, and implements national security strategies. From this standpoint, data interception is typically read according to two different modalities, one concerning citizen and the other regarding non-citizens. When it comes to the foreign services, the exchange of data is not considered to be routine practice. Many scholarly works give the impression that most intelligence services are reluctant to share data, notably due to their commitment to secrecy. When data exchanges do occur, authors argue that this is only happening between national intelligence services that have developed mutual trust in the fight against shared enemies, such as during the World War II and the Cold War. While not taking issue with some aspects of this reading that may be erroneous, it is at the very least far too monolithic.

To challenge this depiction, over the last couple of years, we have developed a Bourdieusian-inspired analysis of the contemporary international by pointing to transnational fields of power, their dynamics, and the dispositions that the actors enact when what is at stake is the management and extraction of data for purposes of constituting watch lists of suspects (Ben Jaffel 2018, Bigo 2016, Bonelli and Ragazzi 2014). Specifying the activities of intelligence services into the general management of unease by security professionals, the idea of a guild of extraction of sensitive information has been proposed to analyze the current composition and roles of the different intelligence services in countries claiming that they are democracies and that they accept the idea of limits to secret, intrusive practices

regarding the whole population, be it regarding their citizens or foreigners (Bigo 2018). As we shall see later in this chapter, our research shows that in this particular area of intelligence, we can observe specific social universes that relate to the objective properties of intelligence services and the manner that they construct intelligence data. It seems that in the case of intelligence practices linked to "global" counter-terrorism—and likely other missions—transnational logics and allegiances are stronger than purely national ones. As intelligence data are constituted by the types of questions that are raised and methods of reasoning that are used, data exchanges are actually more routine between services that belong to different countries but that share the same visions, know-how and practices concerning intelligence objectives than they are between services of the same country that deploy different or even complementary practical know-how. It is therefore necessary to understand the emergence of so-called trans-governmental networks between intelligence agencies, or more accurately transnational guilds that bring together agencies sharing the same specialized visions and whose agents have similar dispositions emanating from their socialization at work. It is therefore the professional habitus that can allow individuals and agencies to overcome national differences. In some cases, loyalties between agencies can be stronger than an institutional attachment to the political leaders of their country. To name just one example from a series of recent cases, in the context are nearly-routine exchanges, it seems that one department from the Bundesnachrichtendienst (BND) entrusted confidential information about the government of Angela Merkel and German politics to their National Security Agency (NSA) allies (Hegemann and Kahl 2016).

There is therefore a transnational intelligence space made up of different groups of national services that cooperate together in the management of digital data and of sensitive information, more generally. This transnational space is not structured and divided according to the national policies of governments—even if they do play a role by way of existing coordination structures. Instead, this space is defined by the types of information that actors seek out, the characteristics of these services, their composition, and their practices. So, what are the relationships between three different universes, which are each defined by different practices of intelligence as an occupation?

In order to clarify our working hypothesis, we have made a first attempt to map the transnational space of intelligence agencies in countries that agree to call themselves democracies and, at the same time, have regional or global foreign policy ambitions. Embracing a methodology that draws from the work of Pierre Bourdieu and international political sociology, we have used multiple correspondence analysis (MCA) to draw this space. As a methodological tool, MCA essentially allows us to mathematically distribute the services in a two-dimensional space, gathering them according to their most significant resemblances and differences. While allowing for a more systematic analysis of qualitative data collected in interviews, MCA is a heuristic tool for identifying groups, which are by no means randomly constituted (Le Roux and Rouanet 2010).

The sheer size of the American services, and their budget, explains their overwhelming engagement in cooperative initiatives and their role as the leaders of

networks that bring together countries from the so-called Global North that engage in activities beyond liberal democracies (Bigo, Bonelli, and Deltombe 2008). Since Edward Snowden's disclosures on the activities of the NSA, the Five Eyes has become the most well-known of such networks. It brings together agencies from the United States, Great Britain, Canada, Australia, and New Zealand that work with satellite, electronic, and internet communications data. This network is no longer limited to a group of five, as now it now includes more than a dozen intelligence agencies or sub-units from countries like France, Germany, Spain, and Sweden, which have all put in place infrastructures for the interception of data passing through submarine and terrestrial cables.[13] Based on the understanding that nowadays nine countries are involved in the Five Eyes Plus network, we have selected 25 agencies from the represented member countries, which all claim to be democracies and also have the ambition of playing a regional or global political role. Categories of intelligence service agencies include counter-terrorism and counter-intelligence services (police and non-police), external intelligence services, and, when they exist, the technical services dedicated to large-scale data interception. For the time being, military intelligence agencies working on military-specific issues have been excluded from this study, although they do sometimes play a role in diplomacy or even the fight against terrorism. Financial intelligence services also at times become involved in the fight against terrorism but have also been left out as their main activities relate to anti-money laundering. Border control agencies have also been excluded.

TEXTBOX 6.1 METHODOLOGICAL DETAILS

For the United States, we have selected the following agencies: National Security Agency (NSA), Central Intelligence Agency (CIA), and the Federal Bureau of Investigation (FBI)[14]; for Canada: Canadian Security Intelligence Service (CSIS) and Communications Security Establishment (CSE); for the United Kingdom: Counter Terrorism Command (CTC), Security Service (MI5), Secret Intelligence Service (SIS or MI6) and Government Communications Headquarters (GCHQ); for New Zealand: New Zealand Security Intelligence Service (NZSIS) and Government Communications Security Bureau (GCSB); for Australia: Australian Signals Directorate (ASD), Australian Security Intelligence Organisation (ASIO), and Australian Secret Intelligence Service (ASIS); for France: Direction générale de la Sécurité extérieure (DGSE), Direction générale de la Sécurité intérieure (DGSI) and Service central du renseignement territorial (SCRT); for Germany: Bundesnachrichtendienst (BND), Bundesamt für Verfassungsschutz (BfV), and Bundeskriminalamt (BKA); for Spain: Centro Nacional de Inteligencia (CNI), Comisaría General de Información (CGI) and Servicio de Información de la Guardia Civil (SIGC); and for Sweden: Försvarets Radioanstalt (FRA) and Säkerhetspolisen (Säpo).

(continued)

(continued)

For each of these 25 services, we analysed the following 9 active variables:

1. The territory of competency, with three modalities (internal, external, internal/external);
2. The legal authority of agents, with two modalities (yes, no);
3. Operational capacities, that is whether the agency has agents on the ground, with two modalities (yes, no);
4. Objectives assigned to the services, with two modalities (the fight against internal threats; the defense of national interest—which includes espionage);
5. The number of personnel, with three modalities (0–1999, 2000–4999, 5000+);
6. Technologies used, with three modalities (human intelligence HUMINT, technological intelligence signals intelligence (SIGINT), and mixed capacities MIXED_TECH);
7. Hierarchical authority, with three modalities (Ministry of the Interior or Ministry of Justice; Ministry of Defense or Ministry of Foreign Affairs; Head of Government);
8. Engagement in counter-intelligence activities, with two modalities (yes, no); and lastly,
9. The ability to conduct clandestine or covert operations, with two modalities (yes, no).

As a result of our MCA analysis, we have produced two graphs. The first graph (Figure 6.1) shows the most significant modalities that structure this transnational intelligence space by exploring the types of capital different services possess as well as their organizational attributes and their institutional objectives. The second graph (Figure 6.2) is based on an analysis of the objective properties of agencies, which allows for the visualization of specific subgroups or universes based on the identification of proximities and distances between the various services.

Axes 1 and 2 explain roughly 64% of associations between our nine categorical variables (respectively 44.03% for Axis 1 and 19.98% for Axis 2). On Axis 1, the most contributing variables on the left side of the graph (negative values) include domestic threats (7.8%), internal territories of competency (7.8%), the absence of clandestine operations (6.1%), counter-espionage activity (4.5%), attachment to the Ministries of the Interior or of Justice (6.1%), judicial authority (4.4%), and human intelligence (4.6%). The contributing variables run in opposition to those that appear on the right side of the graph (positive values), which include national interest (8.5%), internal/external territories of competency (7%), clandestine operations (7.8%), attachment to the Ministries of Defense or of Foreign Affairs (6%), technological intelligence (7%), and mixed human/technological intelligence (1.5%).

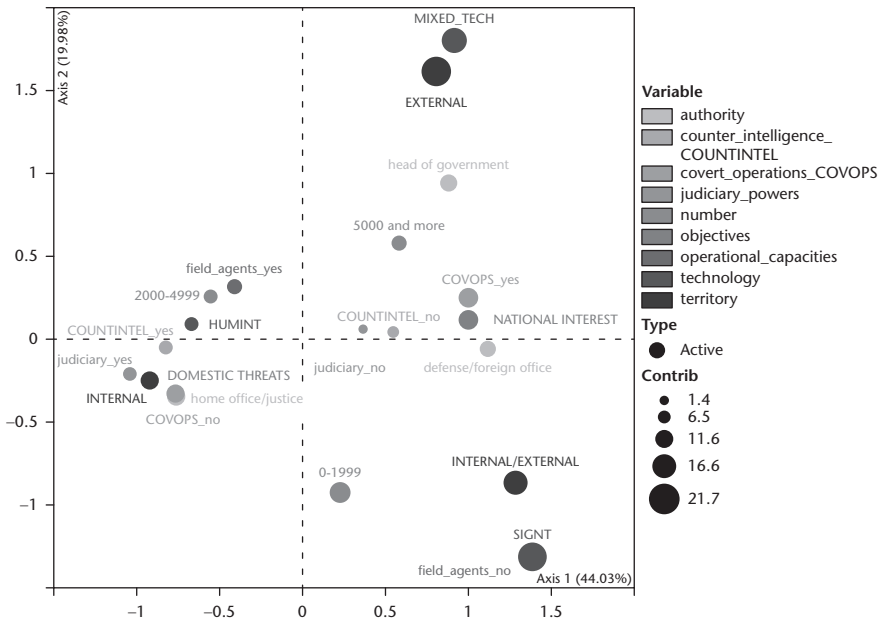


FIGURE 6.1 Most Contributing Variables on Axes 1 and 2 of the MCA

On Axis 2, the most contributing variables on the bottom part of the graph (negative values) include technological intelligence (12.7%), the absence of field agents (12.7%), and internal/external territories of competency (7.5%). In contrast, the most contributing variables on the top part of the graph (positive values) include mixed human/technological intelligence (14.6%), the presence of field agents (4%), and external territories of competency (19.7%).

This distribution of the properties and dispositions that define each of the intelligence service agencies than make it possible to observe the proximities and distances between those institutional actors in space. As shown in Figure 6.2, this exercise allows for the identification of three distinct sets or groups of actors.

The first pole (1), situated on the left in the middle of the table, represents a space of proximity between services that primarily recruit police and prioritize internal security issues. This set of actors is most concerned with internal threats. At times, however, they may develop forms of external actions and cooperation in order to prevent internal threats instigated by actors coming from abroad. They conduct two different categories of missions: (1) political subversion, which includes but is not limited to anti-terrorism; and (2) counter-intelligence. This first subgroup is also defined by their affiliation with Ministries of the Interior or of Justice.

Within this first group, we can identify two subsets: 1) agencies that have at least partial judicial authority (law enforcement agencies) and 2), which do not

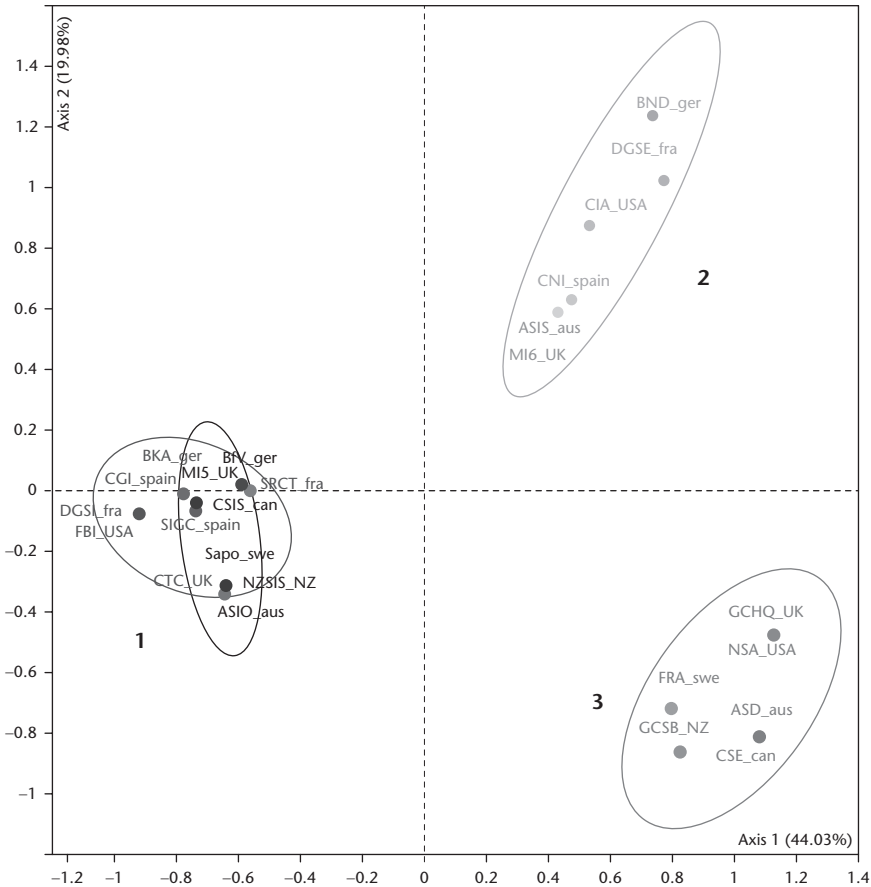


FIGURE 6.2 The Space of Institutional Positions

possess judicial authority. With respect to the first subset, we observe the presence of the German BKA, the Spanish CGI and SIGC, the American FBI, and the French DGSI. Situated in the second subset are the German BfV, the British MI5 and CLC, the Swedish Säpo, the Canadian CSIS, the Australian ASIO, the New Zealand NZSIS, and the French SCRT. Regarding the relationship of these agencies with digital and algorithmic reasonings, qualitative interviews with service agents suggest that they see the evolution of new information and communication technologies (NICTs) as both a constraint and a resource. On the one hand, NICTs are perceived as a constraint because of the flows of information that are now generated by individuals. Indeed, surveillance targets produce far more data today than they did in the past. The volume of available data means that qualitative analyses are a priori not possible. Yet, the amount of digital traces left by individuals is also a valuable resource for intelligence officers. For example,

in addition to telephone calls, SMS messages can be used to demonstrate the frequency of contact between individuals who previously sued other communication channels and who would not have otherwise been noticed. Additionally, digital communication makes it possible to geolocate individuals and to confirm the co-presence of two people in the same place at the same time. Similarly, the consultation and updating of information via Facebook also makes it possible to locate the internet user in question. While no one complains about the wealth of data that is available, these intelligence agencies often pinpoint the question of how to exploit this data as a major obstacle. One possibility is the use of computer software to filter the massive influx of digital data. Some software can also be used to draw graphs of relations based, for example, on correlations between called numbers and localizations in fixed time slots. These graphs can be cross-checked with elements collected using other investigation methods (i.e. witness hearings, interrogations, searches, etc.). It is only as of late, however, that digital data have begun to be used. Since their realization of the potential use they could make of digital data, internal intelligence services have of course adapted their practices to technological evolutions and corresponding societal transitions. However, new technologies have not destabilized the work logics of these agencies. This can be explained by the fact that while the studied agencies have increased staff numbers and strengthened units dedicated to Islamic political violence, they have only very marginally recruited new technological specialists. Therefore, they continue to principally recruit and integrate police offers, agents, and analysts whose skillsets correspond to more traditional intelligence occupational groups. Practical knowledge on how to deal with human sources of information (i.e. informants), shadow suspects, or carry out interrogations continues to be dominant. The most technical tasks can instead be subcontracted to outside parties. For example, in France, geolocalization analyses are entrusted to authorized private companies, which examine the data that has been collected following judicial requisitions and then submit reports to intelligence agents. In this case, digital data actually support and nourish long-term modes of reasoning and institutional practices.

This first subset of actors is clearly distinguished from the second pole (2), which predominately includes agencies that recruit military actors into their ranks, have the operational capacity to missions on external territories, and use espionage techniques. This second subset includes agencies like the CIA and other external service agencies, such as the Spanish CNI, the British MI6, and the Australian ASIS—all situated at the bottom of the eclipse—that rely more on human intelligence. Agencies like the French DGSE and the German BND share similar characteristics to the aforementioned agencies; however, in recent years, these two institutional actors have developed important data interception capabilities within specific departments that are dedicated to the interception of digital data so as to monitor social networks. Yet, they all remain under the general supervision of the Ministry of Defense or of Foreign Affairs. They also tend to see internal intelligence services as potential "clients," which can make specific service requests

to this second subset of actors. Amongst this subgroup, computer-based tools are predominately used to geolocate external targets, to keep in touch with overseas agents, and, occasionally, to drive armed drones in particular operational contexts. When carrying out politically costly operations, big data should thus not create confusion or inaccuracy in hitting targets, meaning that approximation is not really allowed. Digital data is not totally dismissed but is instead only used as a tool of exploration as it is responsibility on the ground, which comes before anything else.

The third pole (3) indicates the emergence of an autonomous subset of SIGINT-Internet agencies. The specificity of this group is that included agencies lack operational agents. Instead, these agencies provide other national intelligence services, both internal and external agencies, with the satellite, terrestrial, and digital data they need. This pole consists almost exclusively of Five Eyes agencies, with the exception of the Swedish FRA, which more or less joined the Five Eyes because of its role in the interception of terrestrial and submarine international cables going to and from Russia. This particular space is at the origins of a new mode of reasoning that delegitimized the effectiveness of traditional forms of intelligence when faced with small, unknown groups. Before the 2000s, Admiral Poindexeter alluded to the development of a global data system identifying targets not based on the pinpointing of individuals already known to intelligence services, but instead through the detection of behavior abnormalities that do not correspond to system logics. Initially called Total Information Awareness (TIA), this mode of reasoning was subsequently referred to as "collect it all" for detecting "weak signal" (i.e. a needle in a haystack), which consists of grouping individuals who did not necessarily know each other together into collectives based on their association with a specific risk profile (Ericson and Haggerty 2006, Harris 2010, Murray 2010). Recalled by its old friends inside the new administration, following 2001, Poindexeter had the means to realize this project of TIA, but the US Senate rejected it at the time for other reasons. Only a small part of it was enforced. Since, however, with its extraordinary capacity in terms of staff and budget, the NSA has once again embarked itself on this journey. But this time, the NSA has brought in the private sector, from data mining software companies and internet providers (i.e. the GAFAM) to telephone companies like Verizon.

Following the Snowden disclosures about the practices and ambitions that guided the NSA, we now know that several intelligence agencies, including the British GCHQ, argued that the potential surveillance of all would never work in operational terms. Instead, digital data collection needed to target small groups so that it could be complemented by human surveillance, judged as more effective. Other services followed the GCHQ in its strategy of recalibration. For example, it appears that services involved in the interception of sensitive information wanted to obtain the necessary legal facilities to be able to undertake large-scale surveillance of potentially-dangerous groups by throwing a "broad net," while simultaneously rejecting an algorithmic mode of reasoning. Interestingly, it is more financial surveillance services or, more recently, services controlling what type of people are authorized to cross borders—that is the newcomers in the intelligence

field—who make claims to being able to handle large amounts of data and persons by using weak signal approaches in order to predict their behaviour. According to them, the ethical-political costs of making false positives are not so important when dealing with suspected persons, they just have to "wait longer" on queuing.

This distribution in a two-dimensional space—which is not random—allows us to group together services from different countries according to the structural proximity of the type of institutional objectives they defend and the know-how they employ. In doing so, this method allows us to visualize cleavages between services from the same country, thereby contrasted to the dominant national-territorial representation of intelligence agencies. As this mapping exercise suggests, the usual discourses on mutual trust only operate between agencies with similar or identical structural positions. The structuration of two of the three poles around the divide between external security and internal security is a rather historically trivial one. However, the emergence of a third pole around SIGINT-Internet agencies may play a considerable role in the destabilization of the status quo if this specific universe begins to successfully impose its definition of intelligence data according to its own practices. Such a reconfiguration would also require the support of politicians, which to a large extent correlates with preventive and predictive political discourses. So, it is essentially in the 2000s that computerization and digital technologies introduced a new arc of tensions into the transnational intelligence space over who produces, exchanges, and analyzes data. Such tensions, however, already existed between intelligence professionals recruited from the military and those recruited from the police. Yet, it is the end of bipolarity that put into question and challenged existing rules and practices of espionage and counterintelligence.

To complement these findings, it would undoubtedly be interesting to know the extent to which politicians and top civil servants are able to impose imperatives of fusion, homogenization, or strict complementarity with regards to the practices of these three poles. Or whether, instead, these poles are autotomizing and create a transnational space of solidarity, complicity. Intelligence agencies, nevertheless, are to some extent subject to the will of political leaders who have control over budgetary and staff allocations. This effect of political control is more visible in countries that have put in place strong coordination structures. For example, the role of the Joint Intelligence Committee (JIC) in the United Kingdom has been to create and reproduce strong cohesions between national intelligence agencies exactly to avoid transnational struggles and alliances. Yet, as shown in the Feinstein report, other countries, and especially the US, have encouraged these transnational games in order to maintain a shroud of opacity around the actions of typically the least supervised external actions. As we have discussed elsewhere, this phenomenon of dynamics interaction means that in some cases the failures of some represent conditions of happiness for others. This is essentially what happened with the failure and controversy around the practices of torture at distance by the CIA and its accomplices[15]. This allowed for the delegitimization of the use of foreign military services in offensive counter-terrorism operations and allows

for less violent means with the NSA's remote action model, which relied on the identification of terrorists within social networks (numeric and non-numeric) and the algorithmic identification of behavior abnormalities. In this case, the failure of a specific subset of external services and the inefficiency of the strategies they used, led to international disproval, which essentially worked to reinforce the decisions to strengthen SIGINT-Internet intelligence services by giving them the resources they had long been asking for.

Concluding remarks

So, what are the results of our Bourdieusian international political sociology-inspired analysis? First, our analysis disproves the existence of a homogenous world (or community) of intelligence wherein all national agencies are complimentary to one another and wherein the boundaries of their missions are clearly defined by the law or by political authority. Our analysis of the constitutive practices of intelligence actors and their meaning-makings of data has destabilized the illusionary idea of the intelligence community as a single world united by common surveillance techniques that are changing the sense of security (and leading globally to speculation). Logics of action cut across and transgress distinctions between the internal and the external, the national and the foreigner. The apparent unity of a national intelligence community in each country must therefore be deconstructed to highlight the relationships that exist between different poles. These relationships are read in terms of how agencies construct intelligence data-suspect for different purposes, how they negotiate amongst themselves as well as with politicians about what approaches should be considered as the most appropriate. These agencies often compete with one another, not only within national fields but also within a transnational space where solidarities—which may still have a hierarchical nature[16]—are made between agencies deploying more or less identical forms of know-how. This reality undermines the dominant and almost exclusive discourse that national security is a source of legitimate suspicion, making it necessary to evoke a narrative on the prevention of attacks and to design global security policies against terrorism. Regarding the former, the more such narratives are questioned, the more it becomes tempting to present these transformations as the result of the emergence of the internet and digital evaluations, even though in fact this shift is the result of political transformations.

Distinctive logics defining cleavages between agencies are not pathologies. Rather, these differences are inherent to the very structure of the intelligence game. Efforts to merge services may not only reinforce inefficiency by reorganizing and destabilizing relations between different types of know-how but may also lead to the creation of hegemonic structures that will impose singular understandings of what data are and for what ends they should be used. This would undermine the plurality of interpretations as well as the richness of debates and discussions. To speak about these differentiated logics is not a return to the image of a tuff war. Instead, our aim is to provide a deeper understanding of the practices of

these agencies, which goes beyond official organizational charts given by political structures or communications agencies. These distinctive logics may also be reproduced within agencies, depending on which recruitment criteria and forms of socialization are privileged. Some intelligence agencies have opted for strong internal homogeneity, trusting only one type of profession or graduates of a single training school, in order to build solidarity. Others have inversely chosen to take on a diverse range of missions and thereby recruit people who have different characteristics in terms of training, gender, violence management practices, and use of numeric technologies. For example, a network engineer and a policeman obviously do not have the same relation to the digital, but they may nonetheless work for the same institution. While all dealing in some way with data, having know-how on how to use data technologies as a user should not be confused with skills required by software designer or by someone who creates profiles based on algorithms. The same goes for the latter, who build populations of target categories, and those who aspire to achieve the same mission, but do so by generating files on precise individuals and organizations, giving great importance to the individual psychologies and trajectories. These different types of intelligence agents all live and work more as analysts than as combatants, which makes them different from those that are deployed "on the ground" to use coercive means in foreign lands. Depending on one's training, the resources at the disposal of agencies and their legitimacy of their actions, resources, and capitals are unevenly distributed amongst actors situated in the intelligence space.

As we have seen in our interviews, it seems that digital techniques are put to use in two ways. Firstly, they can be used in support of the more traditional framework of conjectural reasoning in order to provide necessary evidence for the judiciary. Secondly, they can also be used to impose a preventive and predictive reasoning. The logics and mechanisms of reasoning that are specific to each universe and its actors—be it the police, military, or communications—are therefore to be considered more important than the technologies themselves. In other words, it is not computer technologies that play a role on their own, but rather it is the entry of computer scientists into intelligence circles and the manner in which they frame problems in relation to technology. It is for this reason that the entry of new technologies should not be overestimated (as some interpretations tend to do). Such assumptions tell us little about the effects of technology on practices. For example, actors may continue to use old paper filed while simultaneously mobilizing computer-based tools simply for their cross-referencing speed. Digital technologies may also be employed in the technological regulation of databases and their interoperability. This exercise imposes certain characteristics and criteria on the formatting of data, which is helpful if they are to be exchanged on a regular basis and in large quantities. Digital technologies may come to play a significant role within specialized departments of the military or police services that are dedicated to identification tasks and are supervised by new technical actors. Beyond recurring tensions and disputes precisely on the performativity of data, clashes may also arise between agents with different dispositions. More precisely, in worlds of

intelligence that previously worked primarily with non-digital information and data, the introduction of new techniques—including sophisticated ones—will not necessarily change existing modes of reasoning. It basically takes time for digital technologies professionals to successfully imposer their own interests and professional visions on more traditional actors situated in the intelligence field. However, private companies have played a substantial role in the recruitment of IT specialists, network engineers, data analysts, integration platform software designers, language and coding specialists, cryptologists, and mathematicians tasked with creating or combining algorithms that play on the recognition of weak signals in long series. While the individuals are employed by private actors with the overall ambition of selling products and services, as they increase in number, they begin to more significantly populate a world that previously consisted almost exclusively of police, gendarmes, military, internal intelligence specialists, and external border guards. As a result, these new actors change the rules of the game of these social universes, and in changing rules and habits, they end up changing certain dispositions. This is notably the case in the establishment of good working relations between the world of private contractors and that of public service agents.

Issues and tensions raised by the emergence of this new category of sensitive information professionals have therefore not had a uniform, even impact across different intelligence universes. However, in terms of their transversal impact, this group of professionals has nonetheless attracted the attention of those increasingly supporting preventive and predictive approaches. And together, they have defended the idea that it is only the potentialities and possibilities that digital data bring to the world of intelligence that can satisfy the desires of politicians, the fear of populations, and the interests of security apparatuses in the remote management of populations. Now an essential element in the world of intelligence and surveillance, the capacities of digital techniques nonetheless continue to be debated, all the while linking high politics with the everyday and redefining the way national security is understood.

Notes

1 The terminology used by the services is "raw data" for the data that are generated either by a research they launch or by what is called "captured data", usually by machines or terminals, as a secondary function. For example, cash registers, smartphones, and speedometers serve a main function but may collect data as a secondary task. In informatics, specialists called it "exhaust" data.
2 See note 7 for an explanation of this term.
3 Regarding the notion of field of professionals of sensitive information see Bigo 2018.
4 See in this book the chapter by Ronald J. Diebert and Lou W. Pauly.
5 Symbolically, the attempt to create a distinction between meta-data and data is a way to justify that data are not the property of the internet-user. It justifies exploitation of data and their circulation, compilation, disaggregation and reaggregation outside the knowledge of the individual at the source of the data.
6 On October 19, 2016, the Court of Justice of the European Union (CJEU) decided that the dynamic IP address of a website visitor is "personal data" under Directive 95/46EC (Data Protection Directive) in the hands of a website operator that has the means to compel an internet service provider to identify an individual based on the IP address.

7 The *encomienda* was a forced labor system prevalent in the Spanish Empire, whereby natives were stripped of their property rights. The Spanish Crown gave parcels of land to private individuals that worked in its name, and with that land additional natives who, in theory, worked in exchange for protection and their religious conversion. The relationship between intelligence agencies, the Big Four, and individuals is comparable to the *encomienda* system in so far as internet users are refused ownership to their own data and the work that they do to produce and diffuse them.
8 See Engin Isin and Evelyn Ruppert in this book.
9 The General Data Protection Regulation 2016/679 is a regulation in EU law on data protection and privacy for all individuals within the European Union and the European Economic Area. It also addresses the export of personal data outside the EU and EEA areas. See Elspeth Guild in this book.
10 The term performance has many different meanings in English, which converge with what we insist on. Performance means achievement simultaneously results, outcomes, findings, but also benefit, delivery, and show, spectacle. In some ways the three connotations are simultaneously true.
11 Most debates on intelligence data exchange are based on the *potential* offered by technology rather than the actual practices of intelligence service agents. This is also the case regarding the regulation of technology and road regulation. For example, it is not because a vehicle can travel at a constant speed of 200 kilometers per hour that it is allowed to do so and that the driver does so. The legal system is there to set limits and restrict technical potentialities. Authors like Louise Amoore and Marieke de Goede have sometimes framed practices and potentialities in equal terms, while equating present circumstances with emerging trends. This has led to an overly programmatic view of intelligence services and their intentions. This is demonstrated in the way that newcomers are treated as emblematic of paradigmatic changes, when in fact there are struggles against the transformations brought forth by new actors, as is the case regarding the validity of the accumulation and retention of data alongside the accuracy of predictive algorithms. See notably de Goede 2008 and Amoore 2011.
12 The data derivative comes into being from an amalgam of disaggregated data reaggregated via mobile algorithm-based association rules and visualized in 'real time' as risk map, score or color-coded flag. As explained by Louise Amoore: It is not that derivative forms supersede disciplinary data modes, and indeed among the reaggregated data elements are conventionally collected visa and passport data, but rather that the relation between the elements is itself changed.
13 Some journalists have spoken of the 9-Eyes, with the addition of Sweden, France, Spain and Germany; or even of the 14-Eyes with Belgium, the Netherlands, Italy, Norway and Denmark.
14 Given the range of missions conducted by the FBI—spanning from criminal police work to internal intelligence—we have only taken into account 3,600 agents situated in the Counterterrorism Division.
15 Guild, Elspeth, Didier Bigo, and Mark Gibney, eds. *Extraordinary Rendition: Addressing the Challenges of Accountability*. Routledge, 2018.
16 For example, such a hierarchy is established by the fact that the NSA has more staff at its disposition than all of the European services combined.

Bibliography

Amoore, L. 2011. Data derivatives: On the emergence of a security risk calculus for our times. *Theory, Culture & Society* 28(6), 24–43.
Amoore, L. 2013. *The Politics of Possibility: Risk and Security beyond Probability*. Duke University Press.

Bauman, Z., and D. Lyon. 2013. *Liquid Surveillance: A Conversation*. John Wiley & Sons.
Ben Jaffel, H. 2018. Britain in Europe, Europe in Britain: The Field of Anti-Terrorism Intelligence Cooperation. PhD dissertation, King's College London.
Bigo, D. 2008. "Globalized (In)Security: The field and the Ban-Opticon", in Bigo, D. and Tsoukala, A. (eds.), *Terror, Insecurity and Liberty: Illiberal Practices of Liberal Regimes after 9/11*. Routledge 10–48.
Bigo, D. 2012. "Security, surveillance and democracy" in Ball, K. and Lyon, D. (eds.), *Routledge Handbook of Surveillance Studies*. 277–285. Routledge.
Bigo, D. 2013. "Sécurité maximale et prévention? La matrice du futur antérieur et ses grilles" In Cassin B., Derrière les grilles: sortir du tout évaluation, 111–138. Fayard.
Bigo, D. 2016. Sociology of Transnational Guilds. *International Political Sociology* 10(4), 398–416.
Bigo, D. 2018. "Beyond national security, the emergence of a digital reason of state(s) led by transnational Guilds of Sensitive Information. The case of the Five Eyes Plus Network", in Wagner, B., Kettemann, M. C., and Vieth, K. (eds.), *Research Handbook on Human Rights and Digital Technology*. Edward Elgar.
Bigo, D., L. Bonelli, and T. Deltombe. 2008, Au nom du 11 septembre: Les démocratie à l'épreuve de l'anti-terrorisme, La Découverte.
Bonelli, L., and F. Ragazzi. 2014. Low-tech security: Files, notes, and memos as technologies of anticipation. *Security Dialogue* 45, 476–493.
De Goede, M. 2008. The politics of pre-emption and the war on terror in Europe. *European Journal of International Relations* 14(1), 161–185.
De Goede, M. 2012. *Speculative Security: The Politics of Pursuing Terrorist Monies*. University of Minnesota Press
Ericson, R. V., and K. D. Haggerty. 2006. *The New Politics of Surveillance and Visibility*. University of Toronto Press.
Gill, P., and M. Phythian. 2016. What is intelligence studies? *The International Journal of Intelligence, Security, and Public Affairs*, 18(1), 5–19.
Ginzburg, C. 1980. Morelli, Freud and Sherlock Holmes: Clues and scientific method. *History Workshop* 9, 5–36.
Guild, E., D. Bigo, and M. Gibney, eds. 2018. *Extraordinary Rendition: Addressing the Challenges of Accountability*. Routledge.
Harris, S. 2010. *The Watchers: The Rise of America's Surveillance State*. Penguin.
Hegemann, H. and M. Kahl. 2016. Re-Politisierung der Sicherheit? *ZIB Zeitschrift für Internationale Beziehungen* 23(2), 6–41.
Lehr, P. 2019. *Counter-Terrorism Technologies: A Critical Assessment*. Springer.
Le Roux, B. and H. Rouanet. 2010. *Multiple Correspondence Analysis*. Sage Publications.
McElreath, D. H., M. Graves and C. J. Jensen III. 2017. *Introduction to Intelligence Studies*. Routledge.
Murphy, C. 2016. *Competitive Intelligence: Gathering, Analysing and Putting it to Work*. Routledge.
Murray, N. 2010. Profiling in the age of total information awareness. *Race & Class* 52(2), 3–24.

7

FROM FAKE TO JUNK NEWS

The data politics of online virality

Tommaso Venturini

"Fake news" is a key subject of data politics, but also a tricky one. As this chapter aims to show, the various phenomena signified by this misleading label have little in common, except being opposite to the kind of algorithmic intelligence that most other chapters present as the main concern of data politics. This does not mean that "fake news" is not related to computational analytics or political intentions, but it does mean that this relation is not straightforward.

To discuss this relation, I will go through a three-stage argument. First, I will criticise the notion of "fake news", dismissing the idea that this type of misinformation can be defined by its relationship to truth. Second, I will propose a different definition of this phenomenon based on its circulation rather than of its contents. Third, I will reintroduce the connection to data politics, by describing the economic, communicational, technological, cultural and political dimensions of junk news.

Junk news is not about algorithmic persuasion

The first stage of my argument consists in showing that the "data" and "politics" of "fake news" are not where they are supposed to be. This entails questioning the idea that "fake news" results from sophisticated psy-ops, based on computational techniques processing social media data to distil highly persuasive messages and dispatch them to the most suggestible audiences. This idea has been popular in the debate over Cambridge Analytica (see Venturini & Rogers, 2019). Cambridge Analytica (or CA) is a disreputed marketing firm that, according to its own admissions, maliciously acquired data on several millions of Facebook profiles and used it to push Donald Trump's election. While the first part of this reconstruction is correct, the second is questionable.

In 2014, CA tried to buy data from the earlier "myPersonality project" of the University of Cambridge (Stillwell & Kosinski, 2012). The project was based on

a personality quiz delivered through a Facebook app, which collected the quiz answers and data on the user's activity in the social network. When the negotiations failed (because the researchers refused to lend their data to non-scientific uses), CA commissioned Aleksandr Kogan to replicate the protocol and collect a new batch of data. Kogan created a similar quiz, but introduced two crucial differences: first, he recruited his respondents through the microwork platform Amazon Mechanical Turk; second, it collected data not only on the users taking the quiz, but also on their friends (as allowed by the Facebook's API until 2015). Through this "indirect harvesting", Kogan was able to collect information on millions of Facebook profiles even though fewer than 300,000 people took his quiz.

Besides violating several basic ethics principles, the protocol developed by Kogan made it impossible for CA to actually carry out the sophisticated "psychographic analysis" that the consultancy boasts as its competitive advantage. Being harvested indirectly, 99.5% of the CA records do not contain any psychological information. Though the original research at Cambridge suggested the possibility to infer personality traits from Facebook traces (Kosinski, et al., 2013 and Youyou et al., 2015), it remains unclear whether such inference can yield subtler information than classic marketing. Kogan himself admitted that the standard Facebook's advertisements have better coverage and segmentation. Investigations carried out by *The Guardian* (Cadwalladr, 2018) and Channel 4 (2018) further indicated that CA's services might rely less on algorithmic intelligence than on standard disinformation techniques (eg. defamation, bribery and honey traps).

The Cambridge Analytica affair suggests that computational misinformation might be a marketing myth. But the definition of "fake news" as algorithmic propaganda is also problematic because it presupposes that the goal of misinformation is deception. At close inspection, however, most of the contents that constitute the present upsurge of misinformation does not appear to ask for the "cognitive adherence" of their addresses. Another example will illustrate this claim.

One of the fake contents most circulated during the 2017 French presidential campaign was a story about Emmanuel Macron (later to become the French president) being homosexual and supported by a gay lobby. The most interesting thing about this story was that its falsehood was never in question (Bounegru et al, 2018). While hundreds of websites and social media accounts retransmitted the story, the vast majority explicitly labelled it as false. Apart from the original publication on the Russian information agency Sputnik News, few sources credited the rumour. Most venues cited the story to debunk it and to exhibit the trophy of a French fake news. The "Macron-is-gay" story struck a chord not because people believed it, but because it incarnated the "fake news" imagery: it involved the Russian propaganda; had sexual implications; resonated with rumours about Macron's wedding, etc. While the "Macron is gay" story had little resonance, the "Russian-propaganda-helps-French-online-trolls" story was a resounding success (even Sputnik News soon begun to denounce the story rather than promoting it).

This example reveals how misleading is the label of "fake news". Announcing a "post-truth era" (Keyes, 2004), it presupposes that there was a time in which the

distinction between true and false was unproblematic. Now, if there is a lesson to be learned from half a century of science and technology studies (Hackett et al., 2008 and Jasanoff et al., 1995) is that this separation is never straightforward. This does not mean that true and false are the same, but that their opposition is not binary or static. As STS scholars have shown, a Manichean true/false distinction is not enough to capture the vast spectrum of reliable-yet-not-without-uncertainties status of facts. Even more important: the true/false dichotomy fails to render the way in which enunciations are solidified by the work of all sorts of actors (Latour, 1979 & 2005). Far from being established by sheer force of evidence, facts are built by a complex work of "truth-grounding" (Lynch, 2017).

The notion of "fake news" is misleading because it supposes that malicious pieces of news are manufactured, while reliable ones correspond directly to reality, denying the very essence of journalistic mediation in its efforts to select, combine, translate and present different pieces of information in a news stor (Schudson, 1989 and Tuchman, 1978). The distinction worth making is not between manufactured and unaffected information, but between stories that are supported by a large and honest truth-grounding work and stories that are not.

Many stories labelled as fake news circulate without asking the "cognitive adherence" of those who spread them. Some are openly satirical; others are put out front their ideological biases; others are just titles used to lure readers into clicking. And while some of these contents are meant to trick their readers into believing them, this is rarely their only purpose, instead their objectives "might include acting as monetisable clickbait for viral content pages, doing issue work for grassroots activist groups, grassroots campaigning work for political loyalists and providing humour for entertainment groups (Bounegru et al, in press).

As noted by the director of MIT Center for Civic Media, in a post entitled "Stop saying 'fake news'. It's not helping", the impossibility to define disinformation on the base of its authenticity has turned the notion of "fake news" in

> a vague and ambiguous term that spans everything from false balance (actual news that doesn't deserve our attention), propaganda (weaponized speech designed to support one party over another) and disinformatzya (information designed to sow doubt and increase mistrust in institutions).
>
> *(Zuckerman, 2017)*

Because of its vagueness, the term "fake news" has become a weapon to discredit opposing sources of information (Donald Trump has provided several excellent examples of such use). According to Claire Ward, director of First Draft (an initiative bringing together the main players of journalism and social media):

> The term "fake news" is insufficient and dangerous to use is because it has been appropriated by politicians around the world to describe news organisations whose coverage they find to be problematic. The term "fake news" is being used as a mechanism for clamping down on the free press, and

> serves to undermine trust in media institutions, hoping to create a situation whereby those in power can circumvent the press and reach supporters directly through social media.
>
> *(Wardle & Derakhshan, 2017)*

Junk news as a viral pollution

So, if "fake news" is not about false information, what is it about? As the "Macron-is-gay" example suggests, spread, rather than fakeness, is the birthmark of these contents that should be called "viral news" or more precisely "junk news" for, just as junk food, they are consumed because they are addictive, not because they are appreciated. Shifting the attention from falsity to diffusion does not belittle its sway. Quite the contrary, it suggests that these contents are all the more dangerous because they cannot be defused simply by debunking them. Discussing a widespread fake story about the support that Pope Francis would have offered to Trump, danah boyd observes:

> I can't help but laugh at the irony of folks screaming up and down about fake news and pointing to the story about how the Pope backs Trump . . . From what I can gather, it seems as though liberals were far more likely to spread this story than conservatives. What more could you want if you ran a fake news site whose goal was to make money by getting people to spread misinformation? Getting doubters to click on clickbait is far more profitable than getting believers because they're far more likely to spread the content in an effort to dispel the content.
>
> *(boyd, 2017)*

"Junk news" is dangerous not because it is false, but because it saturates public debate, leaving little space to other discussions, reducing the richness of public debate and preventing more important stories from being heard. Like rumours (Morin, 1969), junk news proliferates by transmission and transformation. In this, it provides a dark illustration of the mechanism through which social phenomena are constructed according to Gabriel Tarde (1890). In his dispute with Durkheim over the fundaments of the nascent sociology, Tarde refused the idea that collective phenomena would be driven by underlying structures. Instead, he claimed that the social consists in the "simple" imitation of individual behaviours and in their progressive alteration (Latour, 2002):

> A social thing [. . .] devolves and passes on, not from the social group collectively to the individual, but rather from one individual [. . .] to another individual, and that, in the passage of one mind into another mind, it is refracted. The sum of these refractions, [. . .] is the entire reality of a social thing at a given moment; a reality which is constantly changing, just like any other reality, through imperceptible nuances.
>
> *(Tarde, 1898)*

Tarde found it difficult to defend its position empirically, because the research methods of his time did not allow to follow collective transmission at the scale and with the sharpness demanded by his argument. This may be possible today thanks to digital traceability (Boullier, 2015 and Latour et al., 2012). Junk news, thereby, is both the dream and the nightmare of Tarde's sociology. The dream, because it offers an opportunity to map the transmission and transformation of collective actions (Venturini, 2018); the nightmare, because it represents an Orwellian degeneration of such mechanisms.

According to Tiziana Terranova (2012), the "psychological economy" imagined by Tarde (1902) has found a dark accomplishment in the contemporary system of the "attention economy". Drawing on Lazzarato (2002) and Stiegler (2008 and 2010), she argues that "modern media enhanced and extended the range and scope of those processes of invention and imitation that for him [Tarde] constituted the essence of economic life" (Terranova, 2012 p. 11), but also 'caused the processes of individuation that connect psychic and social life to be short-circuited, resulting in the destructive hegemony of the short term over the long term' (ibid. p. 12).

I both agree and disagree with this interpretation. While it is useful to turn to Tarde as a reminder that collective virality has not begun with social media, it would be wrong to believe that there is nothing new in the current weave of misinformation. Like most "attention economy" theories, such an interpretation is too generic. According to Crogan and Kinsley (2012) there are three main ways of conceptualising the "attention crisis": as a form of biopower bridling critical thinking (Stigler, 2010); as a push to extend capitalist economy to leisure intellectual activities (Beller, 2009, Lazzarato, 2014; Marazzi, 2008); as a neurological consequence of the exposure to digital technologies (Carr, 2010 and Hayles, 2007). While these arguments are convincing (cognition is certainly a site of political struggle, immaterial labour is indeed threatened by capitalist exploitation and intellectual technologies do affect our neurological capabilities), they are also too broad to account for the specific misinformation treated in this chapter.

Five modes of junk news production

While collective virality is a constant and essential dimension of social existence, "junk information" is a relatively new phenomenon, because only recently virality has become the object of a complex system dedicated to its production and circulation. Such a system is effective (but also difficult to seize by research and regulation) because it brings together developments that are simultaneously:

1. Economic (the establishment of a market for online attention);
2. Communicational (the socialisation of a "prosumer" audience);
3. Technological (the development of behavioural algorithms and spreading bots);
4. Cultural (the development of virality-oriented subcultures);
5. Political (the technique of trolling).

1. The economy of junk news

The economy of virality surfaced in the early 2000s when many Internet companies gave up the hope of selling contents and services and decided to maximise advertising revenues. According to observers (especially Goldhaber, 1997), such evolution derived from the inevitable inversion of "information economy." Because of its abundance, information cannot be the scarce resource driving digital economy. Instead, as observed by Herbert Simon in 1971, its increase gives value to its opposite, namely attention:

> The wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention.
>
> *(Simon, 1971, p. 40)*

Scarcity, however, is not enough. To be sold, attention needs to be "marketized" (Çalişkan & Callon, 2010), which in turn demands a "metrological system" to standardize and quantify the variety of things that we call "attention." Setting up such a system had already proved difficult for broadcasting media (Bourdon & Méadel, 2014) and the solutions found for radio and television could not be applied to the Internet because of the larger number of online sources. Survey based rating systems could assess the visibility of the most top-tier websites, but cannot harness the teeming richness of online offers.

Google solved this problem in the early 2000s with the launch of two services called AdSense (allowing websites owners sell advertising space) and AdWords (allowing web advertisers to buy such a space). The distinctive feature of this network is the automation of its marketplace thanks to two algorithms. As in all advertisement systems, buyers are not interested in attention in general, but in attention on specific matters (if you sell kitchenware you want your ads in cooking blogs rather than in sports forums). In AdWord, this matching is operated by allowing buyers to buy keywords and then displaying the ads on the AdSense websites that PageRank, the algorithm that made Google's fortune as a search engine, (Cardon, 2013; Rieder, 2012), associates to such queries. In this way, PageRank. The second algorithm concerns the auctions through which the price of keywords is established. To allow these auctions to be carried out ceaselessly and with little human intervention, Google implemented a variant of the Vickrey system in which advertisers set their bids independently and the auction is won by highest bidder at the price proposed by the second-highest (Mehta et al. 2007).

This double automation allowed Google to handle micro-transactions unprofitable for traditional advertising agencies and to scale up its network to millions of buyers and sellers (thereby becoming the cornerstone of Google's revenues). Google Ads is thus the precursor (and dominant player) of a series of "advertisement networks" offering to every website, no matter how marginal or flimsy, the possibility to sell its traffic (O'Reilly, 2007).

Another crucial feature of Google Ads is that the value of attention is calculated on the basis of the number of clicks generated by advertisements (unlike previous systems of audience measurement, which could only estimate the number of viewers). This "hit economy" (Rogers, 2002) creates incentives for the thriving of clickbait techniques. "Clickbaiting" is a crucial and yet understudied phenomenon consisting in the proliferation of advertisements with the only objective of being clicked. Clickbaits do seek the sustained attention pursued by newspapers, television, radio or classic websites. They only need to be remarkable enough to pull visitors in for a few seconds. Lowering the barriers of the attention market and merchandising a fleeting attention, the hit economy encouraged the development of a clickbait industry that is responsible for much of the disinformation discussed in this chapter (Graham, 2017, p. 12).

For many junk news producers, the name of the game is not to create catchy stories to generate political effects, but to exploit political interest to clickbait attention. This is famously the case of the teenagers of the Macedonian city of Veles where many highly visited pro-Trump websites and Facebook pages were created (Tynan, 2016 and Subramanian, 2017):

> The young Macedonians who run these sites say they don't care about Donald Trump. They are responding to straightforward economic incentives:. . .they learned the best way to generate traffic is to get their politics stories to spread on Facebook — and the best way to generate shares on Facebook is to publish sensationalist and often false content that caters to Trump supporters. As a result, this strange hub of pro-Trump sites in the former Yugoslav Republic of Macedonia is now playing a significant role in propagating . . . false and misleading content.
>
> *(Silverman & Alexander, 2016)*

2. The communication mechanism of junk news

The fact that Web advertisements are paid by "click" and not by "impression" suggests that online attention is merchandised a form of engagement (albeit a shallow one). Audience studies have long demonstrated that publics are never mere receivers, but always active interpreters (Morley, 1993). Already in 1981, Dallas Smythe observed that the commodity sold by mass media is the "labour power" extracted from their audiences: "the work which audience members perform for the advertiser to whom they have been sold is learning to buy goods and to spend their income accordingly" (p. 243).

Such an "audience power" has become more important as online platforms have learnt to exploit even the lightest engagement of their audiences. In the 1990s, a distinction existed online between "posters", the minority of individuals contributing to the life of digital communities, and "lurkers", the silent majority who just read their discussions (Nonnecke & Preece, 2003). Such distinction has been thinned by the advent of social platforms, which have drastically reduced the effort necessary

to 'act' online through the so-called "social buttons" (Gerlitz & Helmond 2013). Introduced by content aggregators like Reddit and Digg, such buttons can be placed on any webpage and allow visitors to share such page on the aggregator. Thanks to social buttons, recommending a content has become as easy clicking on it.

"Like" and "share" buttons allow embedding in Facebook contents coming from virtually any webpage, making it possible to consume exteral contents without leaving the platform. Removing the need to manually copy and paste URLs, "these buttons facilitate the cross-syndication of web content and . . . introduce a participatory and user-focused approach to recommendation" (Gerlitz & Helmond, 2013, p. 1351). Together with the automatic re-publishing of friends' posts on personal profiles, the one-click-share function dissolves the distinction between lurkers and posters. All Facebook users are posters, for they all contribute to the circulation of contents by simple fact of having friends and liking contents.

The difference between the audience commodity on traditional mass media and on the Internet is that in the latter the users are also content producers: "the users engage in permanent creative activity, communication, community building and content production . . . [This] does not signify a democratization of the media towards participatory systems, but the total commodification of human creativity" (Fuchs, 2009, p. 82).

This does not mean that all uses of social platforms are shallow – as proven by activists all over the world (Gerbaudo, 2012). It means that social platforms rely on the merchandising of a click-and-share engagement that, in turn, encourages the circulation of messages that are not only sticky, but also "spreadable", i.e. designed to be circulated and engaged with. According to Jenkins et al. (2013), spreadability is obtained by "the use of shared fantasies, humor, parody and references, unfinished content, mystery, timely controversy, and rumors" (p. 202) – all elements typical of online misinformation. Even though falseness is not always associated with virality, it is not by chance that many junk news stories are also false. Precisely because they are not meant to obtain a deep cognitive adhesion but to arouse a superficial click-and-share engagement, junk news leverages the bias of fast thinking (Kahneman, 2011). Exaggerated, hyper-partisan stories are highly "spreadable" (Mihailidis & Viotty, 2017) and this explains why the majority of junk news is more similar to satire than to journalism (Horne & Adali, 2017).

3. The technology of junk news

Social media platforms did not just prepare the ground for junk news, they also set up a technological system to nurture it, through a series of techniques to maximise the "audience labour". As noted earlier, one novelty of Google Ads is the billing by clicks (rather than by impressions). The clickthrough rate is also increasingly complemented "conversion rates" quantifying product sales, service registration, but also application downloads or content redistribution (Hoffman & Novak, 2000).

Facebook, for instance, proposes eleven different advertising objectives defined as "what you want people to do when they see your ads" divided into three groups: awareness (brand awareness, reach), consideration (traffic, app installs, engagement, video views, lead generation, messages) and conversion (conversions, catalogue sales, store visits) (Facebook Business, 2018). Google uses its technological network to allow quantifying website actions (through Google Analytics), video views (through YouTube), phone calls (through forwarding numbers), app installs and app actions (through Android) and even offline actions (through customer relationship software).

Furthermore, platforms are not the only players in the tracking economy. Nowadays, most webpages contain software tracing visitors' behaviours and many sell this information to companies that aggregate and resell them (Anthes, 2015 and Crain, 2018). Increasingly, data brokers collect data directly through "third-party cookies". A "cookie" is a file that stores information on users in the memory of their web-browser. Traditionally, cookies were used by websites to collect information about their users and provide a more personalised experience. Lately, however, websites have started to be paid to host cookies belonging to data brokers (Mayers & Mitchell, 2012). Thanks to these "third-party cookies", junk news websites can monetise their traffic even when their visitors do not click on advertisements. Clickbaits sell audiences' attention indirectly, by helping data brokers to collect information on Internet users that will allow to feed them personalised advertisements on other websites.

Online tracking become thus a way to improve the matching between advertisements and audiences. Once again, this mechanism was universalised by Google Ads, which does not simply award spaces to the highest bidders, but "weights" tenders on the basis of their "quality score" (Jansen & Mullen, 2008). The quality score is crucial, because it implies that advertisements with high "quality scores" can win auctions by bidding less than their competitors (Geddes, 2014, p. 215–256). This score measures the match between a specific advertisement and a specific auction and is computed through an undisclosed formula:

> Real-time, auction-specific quality calculations of expected clickthrough rate, ad relevance, and landing page experience, among other factors, are used to calculate Ad Rank at auction time. These factors, which are based on things known only at the time of the auction, can heavily influence the quality of the user's experience.
>
> *(Google Support 2018)*

The expected clickthrough rate (the probability that users will engage with the ad) is the most important of these factors. This makes perfect sense in a system in which revenues are generated only when users actually click on an announcement. Favouring the expected engagement benefits both advertisers and hosting websites (as well as Google, which retains 32% of the paid price) and, crucially for the argument of this chapter, introduces a positive feedback between tracking and engagement.

Through this feedback data collected on online behaviours becomes the basis for promoting the same behaviours. These feedback mechanisms are heavily deployed

in all online platforms and are an integral part of their addictive power. YouTube, for example, measures the time you spend on videos (and the time spent by users with similar viewing habits) to suggest videos that will maximise your viewing time. Likewise, Facebook measures scrolling behaviours to build personal pages that will induce more scrolling – which explains why the platform has reduced the possibility of posting on friends' walls and favoured the automatic composition of "timelines" (Lorenz, 2017). The data used to nourish this behavioural feedback, of course, are also collected outside the platforms through third-party cookies as admitted by Twitter's Growth Director:

> These tailored suggestions are based on accounts followed by other Twitter users and visits to websites in the Twitter ecosystem. I receive visit information when sites have integrated Twitter buttons or widgets, similar to what many other web companies – including LinkedIn, Facebook and YouTube – do when they're integrated into websites.
>
> *(Laraki, 2012)*

The tracking/engagement loop explains the resources invested by media companies in deep learning algorithms (Mackenzie, 2017). Both the "deep" and "learning" nature of these algorithms deserve discussion. "Deep" refers to the plural and layered functioning of the techniques employed in contemporary artificial intelligence (neural networks in particular). Consider the recommendation algorithm recently introduced by YouTube to maximise users' engagement (Covington et al., 2016). Instead of writing a complex equation that would match users and videos, the computation is broken down in dozens of detection blocks, each considering a single feature of the video or the user. The output of each of these elementary blocks becomes the input for a higher level of computation and so forth for several layers (sometime with recursive calls to lower levels). Though each of the blocks is relatively simple, their combination by the machine is extremely difficult to interpret (as acknowledged by Google computer scientists themselves, Olah et al., 2018): "So, when YouTube claims they can't really say why the algorithm does what it does, they probably mean that very literally" (Gielen, 2017, see also Gielen & Rosen, 2016).

Because of their depth, neural networks are then black boxes that can only be validated through their results (which is why these techniques are considered a form of learning). Oftentimes, the ground truth of these comparisons is the human execution of the same operation (e.g., when an image detection algorithm is compared to manual classification). In the case of recommendation algorithms, however, the ground truth is the increase of the engagement, as recognised by YouTube engineers:

> [F]or the final determination of the effectiveness of an algorithm or model, we rely on A/B testing via live experiments. In a live experiment, we can measure subtle changes in click-through rate, watch time, and many other metrics that measure user engagement.
>
> *(Covington et al., 2016, p. 192)*

Deep learning is thus characterized by a "radical behaviourism" (Cardon, 2010): online platforms don't care about why their users engage with them, how engagement is generated, or what engagement even means – the only thing that matters is increasing their measures of clicking, viewing, scrolling. Add to this that, in the last few years, platforms and marketing enterprises have introduced a multitude of "social bots" (Ferrara et al., 2014) performing automatically and on a large scale the same behaviours of users and thus amplifying viral dynamics (Bessi & Ferrara, 2016; Shao et al., 2017).

This is why asking online platforms to implement filters to neutralise junk news is like asking fast-food chains to implement a recipe to reduce junk food consumption. The core of platforms' algorithmic intelligence (and of their business model) lies in the capacity to maximise the virality of online contents, in ways over which their very creators have little control. They may be able to stop blatantly false and mischievous contents, but they will not oppose the very virality that generates their profits.

4. The culture of fake news

While the role of behavioural algorithms should not be overlooked, technology is not the only forces at play in junk news production. People play a crucial role in the system, through the phenomenon of micro-celebrity and the emergence of virality-oriented subcultures. Micro-celebrity refers to the renown obtained through social media by individuals who do not enjoy a high visibility in other arenas (such as sport, show business, economy or politics). While its emergence can be tracked back to reality-TV, micro-celebrity has been encouraged by the practice of social platforms to provide to their users some of the metrics collected for their advertising market.

These "vanity metrics" (Rogers, 2018) tend to capture a superficial kind of attention and measure a celebrity that is ephemeral and shallow – hence the suffix "micro-". This does not mean, however, that micro-celebrity is inconsequential. The idea that every individual has an audience that can be gauged by the same metrics of commercial and political brands (Marwick & boyd, 2011) push many online users to adopt strategies of "personal branding" (Khamis et al., 2017, Marwick, 2015,). This invites many users to curate their online engagement and re-publish viral contents. Vanity metrics mobilise platform users in support of the spreading economy of online media (Hearn & Schoenhoff, 2016). By re-posting contents that they hope will interest their followers, users work to increase their visibility and at the same time contribute to the maintenance of interpersonal communication networks propitious for junk information. Much has been written about how social media allows companies and political leaders to circumvent traditional gatekeepers and address directly their audiences, but much should also be said about the way in which individuals maximising their personal reach amplify commercial (Murphy & Schhram, 2014) and political messages (Vaccari & Valeriani, 2015).

As a cultural phenomenon, virality plays out not only at the individual level, but also at the level of online subcultures. A particularly interesting case is that of

4chan – a popular Internet forum created in 2003 and characterized by two features. First, 4chan encourages its users to post anonymously (the hacker movement "Anonymous" famously got its name from the pseudonym employed by the majority of 4chan posters). Anonymity works as a liberating feature that allows the publication of contents that would be outrageous in most other online and offline venues.

The second feature of 4chan is the way in which it promotes viral ephemerality. As many online forums, 4chan is organised in boards structured as a list of "threads" ranked according to the most recent post. Given the popularity of many 4chan boards, threads are immediately pushed down by new arrivals, unless they are "bumped up" by new comments. Based on a two-week sample, Bernstein et al. (2011), calculated that the median lifespan of a thread in the "random board" (4chan/b/) is 3.9 minutes, with the longest-lived thread lasting 6.2 hours. Though threads life in other boards may be a bit longer (Hagen, 2018), these figures indicate how 4chan has specialised in the kind of ephemeral attention typical of junk news.

Moreover, threads are limited to a maximum of 300 comments, which means that even threads generating lots of comments are eventually closed and deleted. To deal with this mechanism, the 4chan community introduced the practice of summarising popular discussions and re-posting them. This assures that the most popular ideas are constantly distilled and re-posted in a constant swing of fluidification and condensation that closely resemble Tarde's theory of social change.

The combination of a technical incentive encouraging swelling and of a social practice encouraging mutation (both occurring in an anonymous medium) has created an extremely virulent subculture that generated many of the most popular online memetic images (Shifman, 2013) and many of the fake news that polluted the US election (see Tuter et al., forthcoming).

Junk news thus is not only a commercial enterprise, but also a cultural phenomenon. 4chan communities have developed vernacular practices, ideas and expressions. The political board, 4chan/pol/, has been particularly successful in this process of subculture creation, thus becoming the online epicentre of the "alt-right" (Nagle, 2017) and developing a universe of symbolic references that is both rich and opposed to the mainstream culture – in line with Hebdige's (1979) definition of subcultures. 4chan/pol/ posters depict themselves at the population of a fictional country, the Kekistan, with its flag (a green and black version of the Nazi flag); its religion (the cult of the ancient Egyptian god of darkness "Kek"); its symbols (above all the character of "Pepe the Frog"); and its enemies (the "normies" of mainstream culture).

This last element is crucial because much of the Kekistani subculture revolves around the refusal of the "politically correct". According to its members, this refusal justifies the racism, misogyny and extremism of their discourses. Kekistani "shitposters" (as in their self-definition) commonly adopt a mocking attitude and affirm that they do not stand by the ideas they profess but, instead, spread them as jokes and provocations to generate viral effects (the so-called "meme magic") in other platforms such as Reddit (Squirrell, 2017), YouTube (de Keulenaar, 2018) and Twitter (Zannettou et al. 2017).

5. The politics of junk news

It is against this economic, technological and cultural background that pressure groups and governments have seized the political uses of junk news. This use is very different way from classical propaganda (Chomsky, 1991; Jack, 2017) and resembles instead to the campaigns led by "sceptics" against health and environmental regulations. According to Robert Proctor (Proctor & Schiebinger, 2008) and Naomi Oreskes (Oreskes & Conway, 2010), groups of rogue scientists and marketing experts have been financed since the 1950s by industrial groups to counter the mounting evidence on the risks of tobacco smoking and later of acid rain, the ozone hole and climate change. Interestingly, these "merchants of doubts" do not deny these threats directly, but to nurture the doubt emphasising other potential risk causes. In direct confrontations, sceptics would insistently try to displace the discussion to marginal questions or use provocations to make the discussion noisier.

A similar communicational strategy is known under the name of "online trolling" (Bishop, 2014; Schwartz, 2008). Trolls attack online discussions by asking silly questions; insulting other users; blatantly violating the community codes; and, in general, by pushing other users into useless controversies (Lee, 2005; Schachaf & Hara, 2010). Most of the time, trolls are not interested in the contents of the messages they post. Their objective is not to convince their addressees, but merely to provoke them, in a communicational game that bears many similarities to the memetic culture of 4chan (Bergstrom, 2011).

While trolling is classically carried out as ludic activity (Buckels et al., 2014), it has commercial and political equivalents. Firms and marketing companies have long tried to disguise their lobbying activities as forms of native engagement. This form of "sock-puppetry" consists in using false online personae to promote products (a technique known as "shilling" (Stevens et al., 2013; Luca & Zervas, 2016), or to simulate grassroot support (a technique known as "astroturfing," Cho et al., 2011). More recently some companies have started using sock-puppets more aggressively to capture attention through deliberately outrageous messages (Mahdawi, 2015) or by attacking their opponents and disrupting their conversations (Foucart & Horel, 2017). Once again, these forms of corporate trolling are not necessarily meant to spread false information, but to capture or deviate attention.

Political trolling has similar objectives and has been documented by journalistic investigations all over the world. One of the first of evidence that junk news is produced and spread by government agencies for political purposes has been offered by documents in the "Snowden archive," revealing how a unit of the British intelligence agency prepared briefs on how "(1) to inject all sorts of false material onto the internet in order to destroy the reputation of its targets; and (2) to use social sciences and other techniques to manipulate online discourse and activism" (Greenwald, 2014).

Similarly, an investigation by the *New York Times* (Chen, 2015) has documented the setting up of campaigns of viral misinformation in Russia by the infamous

Internet Research Agency. This "troll factory" appears to be less interested in persuading public opinion of pro-Kremlin propaganda than in depriving online debate of all credibility:

> The Internet still remains the one medium where the opposition can reliably get its message out. But their message is now surrounded by so much garbage from trolls that readers can become resistant before the message even gets to them.
>
> *(ibid.)*

The same strategy has been also deployed by the Russian agency as a form of intervention in other countries (MacFarquhar, 2018; Seddon, 2014).

A third example of political trolling is offered by the study of King, Pan and Roberts (2017) of the so-called "50c party members", an army of Internet commentators hired by Chinese authorities to influence public opinion. One of the most interesting findings is that Chinese misinformation campaigns tend to be concentrated in bursts of a few days replicating "the bursts that occur naturally when discussions go viral" (ibid). Interestingly, this claim has not been denied, but in fact acknowledged by the Chinese government:

> The Chinese internet media's largest problem is . . . the amplification of negative and alternative information on Chinese domestic issues caused by opinion formation mechanisms that have been a part of the Internet since it was invented in the US; Chinese society, in the midst of a transformation, does not have the hedging mechanisms to deal with this amplification, so traditional public opinion guidance systems don't seem to be pulling their weight when it comes to overcoming these problems.
>
> *(Appendix B of King, Pan & Roberts, 2017)*

Finally, these communication strategies are amplified by automatic means. In the same way in which "social bots" contribute to the spread of commercial junk news, "political bots" are used for viral warfare (Cook at al., 2014). In their simplest applications, bots are used to artificially increase the metrics of popularity of political leaders; in their most vicious they are employed to "flood" the discussions of opponents (Woolley, 2016):

> During the Arab Spring, online activists were able to provide eyewitness accounts of uprisings in real time. In Syria, protesters used the hashtags #Syria, #Daraa and #Mar15 to appeal for support from a global teatre. . . spambots created by Bahrain company EGHNA were co-opted to create pro-regime accounts. They flooded the hashtags with pro-revolution narratives. This was essentially drowning out the protesters' voices with irrelevant information – such as photography of Syria.
>
> *(Michael, 2017)*

Conclusions

So, junk news is, after all, a form of data politics. Yet, not in the way often imagined. When we consider the power of data, we often conjure some sort of Big Brother dystopia: a centralised organisation monitoring our actions through a technological panopticon and influencing them through cutting-edge persuasion techniques. This is not the case for junk news, which does not draw on the collection of detailed datasets of personal information and does not exploit advanced influencing algorithms. "Fake news" is more prosaically "junk news", for its cycle of production and distribution resembles to the one of junk food.

In a sense, this is even more worrying, because no single organisation (no matter how sophisticated) is as strong as a system with gears in the economy, media, technology, culture and politics. This is the take-away message of this chapter, that junk information is the consequence of a multiplicity of developments emerged in different spheres but directed to the same purpose: to accelerate but also trivialize the dynamics of variation and reproduction that Tarde saw as the two basic forces of collective life.

The creation of a standardised market for online publics and its expansion to the long tail of the Web; the quantification of engagement through metrics of clicking and sharing; the emergence of a flourishing clickbait economy and the diffusion of clickbait techniques to all types of communication; the training of online audiences to contribute to the distribution of junk information; the introduction of third-party cookies and the advent of data brokers; the use of deep learning algorithms amplify the consumption of viral contents; the deployment of armies of social and political bots; the rise of micro-celebrities and the pervasiveness of vanity metrics; the emergence of virality-oriented subcultures in specialised platforms and their spread to mainstream media; and the perfecting of political trolling and discussion hijacking. All these developments are aligned to promote a type of attention and of engagement that (because of their ephemerality and shallowness) are opposite to those necessary for a healthy democratic debate. This alignment is not fortuitous or vaguely inspired by the same zeitgeist, but connected through a series of reinforcing relations that need to be empirically exposed and legally dismantled to slow down the rise of junk misinformation. The multiplication of the appeals to truth and fact-checking miss its target for it refers to a regime of information (that of traditional journalism) that could not be further away from that of junk news. It only by understanding the system of digital virality that we can stand against online misinformation.

References

Anthes, Gary. 2015. "Data Brokers Are Watching You." *Communications of the ACM* 58 (1): 28–30. doi:10.1145/2686740.

Bakir, Vian, and Andrew McStay. 2018. "Fake News and The Economy of Emotions: Problems, Causes, Solutions." *Digital Journalism* 6 (2): 154–75. doi:10.1080/21670811.2017.1345645.

Beller, Jonathan. 2009. The Cinematic Mode of Production: Attention Economy and the Society of the Spectacle. *Learning, Media and Technology*. Vol. 34. doi:10.5860/CHOICE.44-6137.

Bergstrom, Kelly. 2011. "'Don't Feed the Troll': Shutting down Debate about Community Expectations on Reddit.Com." *First Monday* 16 (8): 9–11. doi:10.5210/fm.v16i8.3498.

Bernstein, Michael S, Andrés Monroy-Hernández, Drew Harry, Paul André, Katrina Panovich, and Gregory G Vargas. 2011. "4chan and/b: An Analysis of Anonymity and Ephemerality in a Large Online Community." In *ICWSM*, 50–57.

Bessi, Alessandro, and Emilio Ferrara. 2016. "Social Bots Distort the 2016 U.S. Presidential Election Online Discussion." *First Monday* November. doi:https://doi.org/10.5210/fm.v21i11.7090.

Bishop, Jonathan. 2014. "Representations of 'trolls' in Mass Media Communication: A Review of Media-Texts and Moral Panics Relating to 'Internet Trolling.'" *International Journal of Web Based Communities* 10 (1): 7. doi:10.1504/IJWBC.2014.058384.

Boullier, Dominique. 2015. "Socio Vie et Mort Des Sciences Sociales Avec Le Big Data." *Socio* 4: 19–37.

Bounegru, Liliana, Mette Simonsen Abildgaard, Andreas Birkbak, Jonathan Gray, Mathieu Jacomy, Torben Jensen Elgaard, Anders Koed Madsen, and Anders Kristian Munk. (in press). "Five Provocations about Fake News." *STS Encounters*.

Bounegru, Liliana, Jonathan Gray, Tommaso Venturini, and Michele Mauri. 2018. *A Field Guide to Fake News and Other Information Disorders*. Amsterdam: Public Data Lab. https://ssrn.com/abstract=3024202.

Bourdon, Jérôme, and Cécile Méadel. 2014. Television Audiences across the World: Deconstructing the Ratings Machine. New York: Palgrave Macmillan. doi:10.1057/9781137345103.

boyd, danah. 2017. "Did Media Literacy Backfire?" Points. https://points.datasociety.net/did-media-literacy-backfire-7418c084d88d.

Buckels, Erin E., Paul D. Trapnell, and Delroy L. Paulhus. 2014. "Trolls Just Want to Have Fun." *Personality and Individual Differences* 67: 97–102. doi:10.1016/j.paid.2014.01.016.

Cadwalladr, Carole. 2018. "'I Made Steve Bannon's Psychological Warfare Tool': Meet the Data War Whistleblower." *The Guardian*, March 17. https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump.

Çalişkan, Koray, and Michel Callon. 2010. "Economization, Part 2: A Research Programme for the Study of Markets." *Economy and Society* 39 (1): 1–32. doi:10.1080/03085140903424519.

Cardon, Dominique. 2010. *La Démocratie Sur Internet. Promesses et Limites*. Paris: Seuil.

Cardon, Dominique. 2013. "Dans l'esprit Du PageRank." *Réseaux* 177 (1): 63. doi:10.3917/res.177.0063.

Carr, Nicholas. 2010. *The Shallows: What the Internet Is Doing to Our Brains*. New York: Norton & Company.

Channel 4 News. 2018. "Revealed: Trump's Election Consultants Filmed Saying They Use Bribes and Sex Workers to Entrap Politicians." Channel 4, March 19. https://www.channel4.com/news/cambridge-analytica-revealed-trumps-election-consultants-filmed-saying-they-use-bribes-and-sex-workers-to-entrap-politicians-investigation.

Chen, Adrian. 2015. "The Agency." *The New York Times Magazine*, June 7. https://www.nytimes.com/2015/06/07/magazine/the-agency.html.

Cho, Charles H., Martin L. Martens, Hakkyun Kim, and Michelle Rodrigue. 2011. "Astroturfing Global Warming: It Isn't Always Greener on the Other Side of the Fence." *Journal of Business Ethics* 104 (4): 571–87. doi:10.1007/s10551-011-0950-6.

Chomsky, Noam. 1991. *Media Control: The Spectacular Achievements of Propaganda*. New York: Seven Stories. doi:10.1007/s13398-014-0173-7.2.

Cook, David M., Benjamin Waugh, Maldini Abdipanah, Omid Hashemi, and Shaquille Abdul Rahman. 2014. "Twitter Deception and Influence: Issues of Identity, Slacktivism, and Puppetry." *Journal of Information Warfare* 13 (1): 58–71.

Covington, Paul, Jay Adams, and Emre Sargin. 2016. "Deep Neural Networks for YouTube Recommendations." Proceedings of the 10th ACM Conference on Recommender Systems - RecSys '16, 191–98. doi:10.1145/2959100.2959190.

Crain, Matthew. 2018. "The Limits of Transparency: Data Brokers and Commodification." *New Media and Society* 20 (1): 88–104. doi:10.1177/1461444816657096.

Crogan, Patrick, and Samuel Kinsley. 2012. *Paying Attention: Toward a Critique of the Attention Economy*. Lebanon, NH: Dartmouth College Press.

Daniel Kahneman. 2011. *Thinking, Fast and Slow*. New York. Farrar, Straus and Giroux.

de Keulenaar, Emillie V. 2018. "The Rise and Fall of Kekistan: A Story of Idiomatic Animus as Told Through Youtube's Related Videos." *Open Intelligence Lab*, April 6.

Facebook Business. 2018. "About Advertising Objectives." Advertiser Help Center. https://www.facebook.com/business/help/517257078367892.

Ferrara, Emilio, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2014. "The Rise of Social Bots." *Communications of the ACM* 59 (7): 96–104. doi:10.1145/2818717.

Foucart, Stéphane, and Stéphane Horel. 2017. "Monsanto Papers : La Guerre Du Géant Des Pesticides Contre La Science." *Le Monde*, May 5. https://abonnes.lemonde.fr/planete/article/2017/06/01/monsanto-operation-intoxication_5136915_3244.html.

Fuchs, Christian. 2009. "Information and Communication Technologies and Society." *European Journal of Communication* 24 (1): 69–87. doi:10.1177/0267323108098947.

Geddes, Brad. 2014. *Advanced Google AdWords* (3rd Edition). Hoboken: John Wiley & Sons.

Gerbaudo, Paolo. 2012. *Tweets and the Streets: Social Media and Contemporary Activism*. London: Pluto Books.

Gerlitz, Carolin, and Anne Helmond. 2013. "The like Economy: Social Buttons and the Data-Intensive Web." *New Media and Society* 15 (8): 1348–65. doi:10.1177/1461444812472322.

Gielen, Matt. 2017. "Reverse Engineering The YouTube Algorithm: Part II," February. https://www.tubefilter.com/2017/02/16/youtube-algorithm-reverse-engineering-part-ii/.

Gielen, Matt, and Jeremy Rosen. 2016. "Reverse Engineering the YouTube Algorithm: Part I." *Tubefilter*, June 6. http://www.tubefilter.com/2016/06/23/reverse-engineering-youtube-algorithm/.

Glenn Greenwald. 2014. "How Covert Agents Infiltrate the Internet to Manipulate, Deceive, and Destroy Reputations." *The Intercept*, August 20. https://www.theguardian.com/commentisfree/2015/aug/20/advertising-trolls-marketing.

Google Support. 2018. "Things You Should Know about Ads Quality." AdWords Help. https://web.archive.org/web/20180425195711/https://support.google.com/adwords/answer/156066.

Graham, Richard. 2017. "Google and Advertising: Digital Capitalism in the Context of Post-Fordism, the Reification of Language, and the Rise of Fake News." *Palgrave Communications* 3 (1): 45. doi:10.1057/s41599-017-0021-4.

Hackett, Edward J., Olga Amsterdamska, Michael Lynch, Judy Wajcman, and Sergio Sismondo. 2008. *The Handbook of Science and Technology Studies*. Cambridge, MA: MIT Press. http://medcontent.metapress.com/index/A65RM03P4874243N.pdf%5Cnhttp://eprints.lse.ac.uk/28629/.

Hagen, Sal. 2018. "Rendering Legible the Ephemerality of 4chan/Pol/." *Open Intelligence Lab*, April 12. oilab.eu/rendering-legible-the-ephemerality-of-4chanpol/.

Hebdige, Dick. 1979. *Subcultures: The Meaning of Style. Queen's Quarterly*. Vol. 89. London: Methuen. https://ezproxy.bibl.ulaval.ca/login?url=http://search.proquest.com/docview/61194747?accountid=12008%5Cnhttp://sfx.bibl.ulaval.ca:9003/sfx_local??url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&genre=unknown&sid=ProQ:ProQ:socabsshell&atitle=S.

Hine, Gabriel Emile, Jeremiah Onaolapo, Emiliano De Cristofaro, Nicolas Kourtellis, Ilias Leontiadis, Riginos Samaras, Gianluca Stringhini, and Jeremy Blackburn. 2016. "Kek, Cucks, and God Emperor Trump: A Measurement Study of 4chan's Politically Incorrect Forum and Its Effects on the Web," *ICWSM*: 92–101. http://arxiv.org/abs/1610.03452.

Hoffman, Donna L., and Thomas P. Novak. 2000. "Advertising Pricing Models for the World Wide Web." *Internet Publishing and Beyond: The . . . and beyond: The . . .* : 1–22. http://elabresearch.ucr.edu/blog/uploads/papers/Advertising Pricing Models for the World Wide Web [Hoffman and Novak - 2000].pdf.

Horne, Benjamin D., and Sibel Adali. 2017. "This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News." http://arxiv.org/abs/1703.09398.

Jack, Caroline. 2017. "What's Propaganda Got to Do With It?" *Points*. https://points.datasociety.net/whats-propaganda-got-to-do-with-it-5b88d78c3282.

Jansen, Bernard J., and Tracy Mullen. 2008. "Sponsored Search: An Overview of the Concept, History, and Technology." *International Journal of Electronic Business* 6 (2): 114–31. doi:10.1504/IJEB.2008.018068.

Jasanoff, Sheila, Gerard E. Markle, James C. Peterson, and Trevor Pinch. 1995. *Handbook of Science and Technology Studies*. Thousand Oaks, CA: Sage. http://www.amazon.com/Handbook-Science-Technology-Professor-Jasanoff/dp/0761924981.

Jenkins, Henry, Sam Ford, and Joshua Benjamin Green. 2013. *Spreadable Media*. New York: New York University Press. doi:10.1017/CBO9781107415324.004.

Keyes, Ralph. 2004. *The Post-Truth Era: Dishonesty and Deception in Contemporary Life*. New York: St. Martin's Press.

Khamis, Susie, Lawrence Ang, and Raymond Welling. 2017. "Self-Branding, 'Micro-Celebrity' and the Rise of Social Media Influencers." *Celebrity Studies* 8 (2): 191–208. doi:10.1080/19392397.2016.1218292.

King, Gary, Jennifer Pan, and Margaret E. Roberts. 2017. "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument." *American Political Science Review* 111 (03): 484–501. doi:10.1017/S0003055417000144.

Kosinski, Michal, David Stillwell, and Thore Graepel. 2013. "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior." *Proceedings of the National Academy of Sciences* 110 (15): 5802–5. doi:10.1073/pnas.1218772110.

Laraki, Othman. 2012. "New Tailored Suggestions for You to Follow on Twitter." Twitter Official Blog. https://blog.twitter.com/official/en_us/a/2012/new-tailored-suggestions-for-you-to-follow-on-twitter.html.

Latour, Bruno. 2002. "Gabriel Tarde and the End of the Social." In *The Social in Question. New Bearings in the History and the Social Sciences*, edited by Patrick Joyce, 117–32. London: Routledge.

Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford: University Press.

Latour, Bruno, Pablo Jensen, Tommaso Venturini, Sébastian Grauwin, and Dominique Boullier. 2012. "'The Whole Is Always Smaller than Its Parts': A Digital Test of Gabriel Tardes' Monads." *The British Journal of Sociology* 63 (4): 590–615. doi:10.1111/j.1468-4446.2012.01428.x.

Latour, Bruno, and Steve Woolgar. 1979. *Laboratory Life: The Construction of Scientific Facts*. Los Angeles, CA: Sage.

Lazzarato, Maurizio. 2014. *Signs and Machines: Capitalism and the Production of Subjectivity*. Cambridge, MA: MIT Press.

Lazzarato, Maurizio. 2002. *Puissances de l'invention: La Psychologie Économique de Gabriel Tarde Contre l'économie Politique*. Paris: Les empêcheurs de penser en rond.

Lee, Hangwoo. 2005. "Behavioral Strategies for Dealing with Flaming in an Online Forum." *The Sociological Quarterly* 46 (2): 385–403. http://www.jstor.org/stable/4120995%5Cnhttp://www.jstor.org/stable/4120995?seq=1&cid=pdf-reference#references_tab_contents%5Cnhttp://about.jstor.org/terms.

Lorenz, By Taylor. 2017. "The Facebook Wall Is Dead—and Facebook Is Struggling to Get Personal Again." Mic.Com, May 10. https://mic.com/articles/176599/the-facebook-wall-is-dead-social-network-struggles-to-get-personal-again#.CL3KRSusd.

Luca, Michael, and Georgios Zervas. 2016. "Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud." *Management Science* 62 (12): 3412–27. doi:10.1287/mnsc.2015.2304.

Lynch, Michael. 2017. "Post-Truth, Alt-Facts, and Asymmetric Controversies." First 100 Days. http://first100days.stsprogram.org/2017/02/06/post-truth-alt-facts-and-asymmetric-controversies-part-i/.

MacFarquhar, Neil. 2018. "Inside the Russian Troll Factory: Zombies and a Breakneck Pace." *The New York Times*, February 18.

Mackenzie, Adrian. 2017. *Machine Learners: Archaeology of a Data Practice*. Cambridge, MA: MIT Press.

Mahdawi, Arwa. 2015. "Don't Feed the Advertising Trolls." *The Guardian*, August 20. https://www.theguardian.com/commentisfree/2015/aug/20/advertising-trolls-marketing.

Marazzi, Christian. 2008. *Capital and Language*. Los Angeles, CA: Semiotext(e). http://books.google.com/books?id=WO4dAQAAIAAJ.

Marwick, Alice E. 2015. "Instafame: Luxury Selfies in the Attention Economy." *Public Culture* 27 (1 75): 137–60. doi:10.1215/08992363-2798379.

Marwick, Alice, and danah boyd. 2011. "To See and Be Seen: Celebrity Practice on Twitter." *Convergence* 17 (2): 139–58. doi:10.1177/1354856510394539.

Mayer, Jonathan R., and John C. Mitchell. 2012. "Third-Party Web Tracking: Policy and Technology." *Proceedings – IEEE Symposium on Security and Privacy*, 413–27. doi:10.1109/SP.2012.47.

Mehta, A., A. Saberi, U. Vazirani, and V. Vazirani. 2007. "AdWords and Generalized On-Line Matching." 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05) V (August): 264–73. doi:10.1109/SFCS.2005.12.

Michael, Katina. 2017. "Bots without Borders: How Anonymous Accounts Hijack Political Debate." *The Conversation*, January 24. https://theconversation.com/bots-without-borders-how-anonymous-accounts-hijack-political-debate-70347.

Mihailidis, Paul, and Samantha Viotty. 2017. "Spreadable Spectacle in Digital Culture: Civic Expression, Fake News, and the Role of Media Literacies in 'Post-Fact ' Society." *American Behavioral Scientist* 61 (4): 441–54. doi:10.1177/0002764217701217.

Morin, Edgar. 1969. *La Rumeur d'Orléans*. Paris: Seuil.

Morley, David. 1993. "Active Audience Theory: Pendulums and Pitfalls." *Journal of Communication* 43 (4): 13–19. doi:10.1111/j.1460-2466.1993.tb01299.x.

Murphy, T. & Schhram, R. (2014). What is Worth? The Value Chasm Between Brand and Influencers. *Journal of Brand Strategies*, 3(1), pp. 31-40.

Nagle, Angela. 2017. *Kill All Normies*. Winchester: Zero Books.

Nonnecke, Blair, and Jenny Preece. 2003. "Silent Participants: Getting to Know Lurkers Better." In *From Usenet to CoWebs*, edited by Christopher Lueg and Danyel Fisher, 110–32. doi:10.1007/978-1-4471-0057-7_6.

O'Reilly, Tim. 2007. "What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software." *International Journal of Digital Economics* 65 (March): 17–37.

Olah, Chris, Arvind Satyanarayan, Ian Johnson, Shan Carter, Ludwig Schubert, Katherine Ye, and Alexander Mordvintsev. 2018. "The Building Blocks of Interpretability." *Distill*, March 6. doi:10.23915/distill.00010.

Oreskes, Naomi, and Erik M Conway. 2010. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. London: Bloomsbury Press.

Proctor, Robert. 2011. *Golden Holocaust: Origins of the Cigarette Catastrophe and the Case for Abolition*. Berkeley: University of California Press.

Proctor, Robert, and Londa Schiebinger. 2008. *Agnotology: The Making and Unmaking of Ignorance*. Stanford: Stanford University Press.

Rieder, Bernhard. 2012. "What Is in PageRank? A Historical and Conceptual Investigation of a Recursive Status Index." *Computational Culture: A Journal of Software Studies*, 1–28. http://computationalculture.net/article/what_is_in_pagerank.

Rogers, Richard. 2002. "Operating Issue Networks on the Web." *Science as Culture* 11 (2): 191–213. doi:10.1080/09505430220137243.

Rogers, Richard. 2018. "Otherwise Engaged: Social Media from Vanity Metrics to Critical Analytics." *International Journal of Communication* 12 (732942): 450–72.

Sanjuán, Rafael, and Pilar Domingo-Calap. 2016. "Mechanisms of Viral Mutation." *Cellular and Molecular Life Sciences* 73 (23): 4433–48. doi:10.1007/s00018-016-2299–96.

Schachaf, P, and N Hara. 2010. "Beyond Vandalism: Trolls in Wikipedia." *Journal of Information Science* 36 (3): 357–70.

Schudson, Michael. 1989. "The Sociology of News Production." *Media, Culture & Society* 11 (3): 263–82. doi:10.1177/016344389011003002.

Schwartz, Mattathias. 2008. "The Trolls Among Us." *The New York Times*, August 3. http://www.nytimes.com/2008/08/03/magazine/03trolls-t.html.

Seddon, Max. 2014. "Documents Show How Russia's Troll Army Hit America." Buzzfeed, June 2. http://www.buzzfeed.com/maxseddon/documents-show-how-russias-troll-army-hit-america.

Shao, Chengcheng, Giovanni Luca Ciampaglia, Onur Varol, Kaicheng Yang, Alessandro Flammini, and Filippo Menczer. 2017. "The Spread of Low-Credibility Content by Social Bots." https://arxiv.org/pdf/1707.07592.pdf.

Shifman, Limor. 2013. *Memes in Digital Culture*. Cambridge, MA: MIT Press.
Silverman, Craig, and Lawrence Alexander. 2016. "How Teens in The Balkans AreDuping Trump Supporters With Fake News." BuzzFeedNews, November. https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo?utm_term=.cro1E9mye#.bemMKQqmV.
Simon, H.A. 1971. "Designing Organizations for an Information Rich World." In *Computers, Communications, and the Public Interest*, edited by M Greenberger, 37–72. Baltimore, MD: Johns Hopkins Press.
Smythe, D. W. 1981. "On the Audience Commodity and Its Work." *Media and Cultural Studies* 2006: 230–256. doi:10.1177/026858094009002003.
Squirrell, Tim. 2017. "Linguistic Data Analysis of 3 Billion Reddit Comments Shows the Alt-Right Is Getting Stronger." Quartz, August. https://qz.com/1056319/what-is-the-alt-right-a-linguistic-data-analysis-of-3-billion-reddit-comments-shows-a-disparate-group-that-is-quickly-uniting/.
Stevens, B. J., J. Harlan, and D. B. Scibelli. 2013. "The Art of Shill: Internet Product Consumption for Savvy Consumers." *Issues in Information Systems* 14 (1): 311–14.
Stiegler, Bernard. 2008. "Within the Limits of Capitalism, Economizing Means Taking Care." Ars Industrialis. http://www.arsindustrialis.org/node/2922.
Stiegler, Bernard. 2010a. *Taking Care of Youth and the Generations*. Stanford, CA: University Press. doi:2009025390.
Stiegler, Bernard. 2010b. *For a New Critique of Political Economy*. Cambridge: Polity Press.
Stillwell, David J, and Michal Kosinski. 2012. "MyPersonality Project: Example of Successful Utilization of Online Social Networks for Large-Scale Social Research." The Psychometric Centre, University of Cambridge.
Subramanian, Samanth. 2017. "Welcome to Macedonia, Fake News Factory to the World." *Wired*. February 15. https://www.wired.com/2017/02/veles-macedonia-fake-news/.
Tarde, Gabriel. 1890. *Les Lois de l'imitation*. Paris: Félix Alcan.
Tarde, Gabriel. 1898. "Les Deux Éléments de La Sociologie." *Études de Psychologie Sociale*. Paris: Giard et Brière.
Tarde, Gabriel. 1902. *Psychologie Économique*. Paris: Félix Alcan.
Terranova, Tiziana. 2012. "Attention, Economy and the Brain." *Culture Machine* 13: 1–19.
Tuchman, Gaye. 1978. *Making News: A Study in the Construction of Reality*. New York: The Free Press.
Tynan, Dan. 2016. "How Facebook Powers Money Machines for Obscure Political 'news' Sites." *The Guardian*, August 24. https://www.theguardian.com/technology/2016/aug/24/facebook-clickbait-political-news-sites-us-election-trump.
Vaccari, Cristian, and Augusto Valeriani. 2015. "Follow the Leader! Direct and Indirect Flows of Political Communication during the 2013 Italian General Election Campaign." *New Media and Society* 17 (7): 1025–42. doi:10.1177/1461444813511038.
Venturini, Tommaso. 2018. "Sur l'étude Des Sujets Populaires Ou Les Confessions d'un Spécialiste Des Fausses Nouvelles." In *Les Fausses Nouvelles: Nouveaux Visages, Nouveaux Défis*, edited by Pierre Trudel. Montreal: Presses de l'Université Laval.
Venturini, Tommaso, Mathieu Jacomy, Axel Meunier, and Bruno Latour. 2017. "An Unexpected Journey: A Few Lessons from Sciences Po Médialab's Experience." *Big Data & Society* 4 (2): 205395171772094. doi:10.1177/2053951717720949.
Venturini, Tommaso, and Richard Rogers. Forthcoming. "The Cambridge Analytica Affair, the Crisis of API-Based Research and How This May Be a Good Thing."
Wardle, Claire, and Hossein Derakhshan. 2017. "Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking (Report to the Council of Europe)."

Woolley, Samuel C. 2016. "Automating Power: Social Bot Interference in Global Politics." *First Monday* 21 (4): 1–11. doi:10.5210/fm.v21i4.6161.

Youyou, Wu, Michal Kosinski, and David Stillwell. 2015. "Computer-Based Personality Judgments Are More Accurate than Those Made by Humans." *Proceedings of the National Academy of Sciences* 112 (4): 1036–40. doi:10.1073/pnas.1418680112.

Zannettou, Savvas, Tristan Caulfield, Emiliano De Cristofaro, Nicolas Kourtellis, Ilias Leontiadis, Michael Sirivianos, Gianluca Stringhini, and Jeremy Blackburn. 2017. "The Web Centipede: Understanding How Web Communities Influence Each Other Through the Lens of Mainstream and Alternative News Sources." doi:10.1145/3131365.3131390.

Zuckerman, Ethan. 2017. "Stop Saying 'Fake News'. It's Not Helping." www.Ethanzuckerman.Com/Blog/2017/01/30/Stop-Saying-Fake-News-Its-Not-Helping/, January 30.

8

SEEING LIKE BIG TECH

Security assemblages, technology, and the future of state bureaucracy

Félix Tréguer

In June 1831, a Frenchman, Alexandre Ferrier, sought to create the first privately-held optical telegraph line between Calais, in Northern France, and London. Ferrier was an adventurous entrepreneur who did not back away from bold ideas. The privately-owned telecommunications infrastructure he set out to build would not only serve the interests of French industry barons willing to track stock prices in the financial capital of the world; it could also be of use for the diplomatic communications of the French government. The whole plan was risky but, after all, there was no law sanctioning the monopoly of the French state over telegraph networks.

To be on the safe side, Ferrier thought it was best to ask the government for an explicit authorisation. But Casimir Périer, then head of the French government, was hesitant as to what his answer should be. Yes, Ferrier's proposal was unusual but, after all, many political and business elites agreed that the telegraph could be a boon for the emerging industrial revolution (Flichy 2009). Could the government seriously consider meet that demand while keeping its monopoly over the telegraph? Many thought not.

Alphonse Foy – the man Périer turned to in order to make up his mind – had an entirely different view on the matter. As the newly-appointed Director of the Telegraph Service at the Ministry of Interior, Foy wrote a letter that offered a more-than-tepid response to Ferrier's project. "Mr. Ferrier's request is entirely inadmissible," he wrote (Charbon 1991, 12). As a servant of the government, Foy was appalled by the notion that the French state could lose its monopoly over telecommunications infrastructures. As Foy argued, "the existence of this telegraph communication would necessarily harm the present privilege of the government to be the first instructed of all-important news." The government had to be the first one to see and learn about what was going on. But the fundamental belief expressed by Foy's blunt refusal was that a privatised telecommunications infrastructure was a challenge that the modern state simply could not handle. All the techniques of

power – like surveillance and censorship – institutionalised since the 16th century in partnership with private actors to control the subversive effect of the printing press and of postal networks while allowing them to serve the interests of both the state and early capitalism would come crumbling down.

So over the next six years, the French administration worked on a plan to retain its monopoly over the deployment and exploitation of telegraph networks, while starting to open its use to the general public. In 1837, the Minister of the Interior, Adrien de Gasparin, appeared before the Parliament to defend a new law designed to put that plan into effect. The goal was to criminalise every transmission that was not authorised by the government and not sent over public telegraph lines.

But in many respects, Foy and Gasparin fought a rear-guard battle. As it had already done with older communication technologies, the state would soon move to a lighter-handed approach. From the 1840s on, the development of the electrical telegraph accelerated national and transnational communications flows, as the industrial revolution spurred the demand for coordination and communications (Beniger 1986). By the end of the 19th century, private corporations did not only make extensive use of the telegraph and of the new communication technology of the time, the telephone; they also played a growing role in the construction and management of national and international infrastructures to serve the needs of states and globalising market actors (Barty-King 1980; Headrick 2012).

Fast-forward 150 years. Neoliberal policies launched in the 1980s have completed the dismantling of public and private monopolies over telecommunications networks, and digital technologies have profoundly intensified data flows. Once again, states have found ways to transpose their traditional techniques of power to digital communications (Galloway 2004; Goldsmith and Wu 2006). When they do – whether it is to engage in surveillance, censorship or propaganda – they almost always do so in interaction with companies who manage parts of the multi-layered architecture of the Internet.

In the process, actors who occupy key positions in the state's intelligence and now law enforcement agencies have to constantly negotiate alliances with these private actors. Today, that means not only dealing with large firms in the media and telecom sectors whose relations to the security field can be traced back from the 16th and 19th centuries, but also with tech firms that have come to dominate the digital economy – mostly US-based multinational online service providers, software producers and online platforms like Google, Apple, Facebook, Amazon and Microsoft.

In other words, "Big Tech" joins "Big Media" and "Big Telco" as yet another oligopoly commanding over the all-important communication industries, and more generally the global economy. Today, Apple, Amazon, Google/Alphabet, Microsoft, and Facebook have acquired the highest market valuations globally (Statista 2018). Tech is now the largest sector in global capitalisation, amounting to 3,582$bn, before financials (3,532$bn), consumer goods (2,660$bn), healthcare (2,300$bn), oil and gas (1,411$bn) or telecommunications (859$bn) (PwC 2017). Having championed new models based on the algorithmic regulation of online communications as well as regimes of surveillance relying on the

systematic collection and analysis of behavioral data (Fuchs and Trottier 2015), these resourceful companies hold an irresistible appeal for security professionals tasked with controlling communication flows.

This, in turn, is leading to a historic shift in the public-private assemblages regulating communication networks, which actually points to a much wider trend in modern state power. In *Seeing Like a State* (1998), James C. Scott has shown how modern statecraft was built by ensuring legibility through measures, metrics and other "state simplifications" aimed at representing and acting upon both the natural and social worlds. In the age of Big Data, the techniques mastered by Big Tech are now seen as crucial to make the digitised world legible and governable. Faced with swelling data stocks and flows, the state needs to see like Big Tech. This gives way to a negotiation process aimed at co-opting its infrastructures and its data-processing techniques. Big Data governmentality hence spreads throughout the security field and beyond, across state bureaucracies.

Shifting public-private assemblages in the security field

To make sense of these ongoing negotiations, it is useful to start with the concept of security assemblages. With the rise of security privatisation in the context of neoliberal economics, the public-private category has become a key theme in security studies, with research on topics ranging from private security guards to the role of private companies in the logistics of military forces or subcontractors in the intelligence field (Abrahamsen and Leander 2015; Williams 2010). But, against those insisting that the increasing role of private corporations in security is one more evidence of the weakening of traditional state sovereignty, critical security scholars like Abrahamsen and Williams (2010, 23) have instead argued that "privatisation is not a challenge to prevailing structures of authority, but is embedded in, and inseparable from, transformations in governance." Seeking to lay new theoretical foundations for understanding how private security is historically and socially constituted, the authors have introduced the concept of "global security assemblages":

> [Global security assemblages are] transnational structures and networks in which a range of different actors and normativities interact, cooperate and compete to produce new institutions, practices and forms of deterritorialized security governance.
>
> *(p. 90)*

Against those authors who ground the concept of assemblage in the philosophy of Gilles Deleuze (e.g. Haggerty and Ericson 2000), Abrahamsen and Williams instead anchor it in the Bourdieusian concept of "fields" to highlight the evolution of material, symbolic and cultural forms of capital within the security field. Their approach also builds on Saskia Sassen's notion of "disassembly" to highlight the fact that an assemblage is actually a process whereby some components of states are configured in new power structures, as formerly public functions are transferred

to the private sector. "Security assemblage," then, refer to the way security governance is increasingly achieved through fluctuating arrangements of networks of state, corporate and other voluntary actors, which together form "knots of statelike power" (Harcourt 2015).

That the tech industry may play an important role in security assemblages may not be surprising. After all, from Charles Babbage's proposal of an Analytical Engine to Alan Turing's Enigma, the genealogy of computers clearly shows an immediate connection between the development of these technologies, and the needs of modern bureaucracies – whether public or private. Data processing tools associated with statistical work and calculation have historically played a key role in the modern state power (Agar 2003; Desrosières 2002). They have also long been a cornerstone of the military-industrial complex, as evidenced for instance by scholarship on the role of IBM in the Holocaust (Black 2012) or inquiries on the history of the Silicon Valley and its intimate relationship with the US military (Bellamy Foster and McChesney 2014; Edwards 1996; Harris 2014; Lécuyer 2007; Levine 2018; Nesbit 2017).

But what makes ongoing negotiations between Big Tech and the security field particularly interesting is that this oligopoly also originates from a corporate culture marked by a "counter-culture libertarianism" (Barbrook and Cameron 1995) – one that has deep historical roots (Turner 2006). As a consequence, from the point of view of many stakeholders, these organisations first appeared as relative outsiders to the security field. As the process of hybridisation between the state and the new masters of communication industries unfolds and has yet to stabilise, the security field intersects with other social fields that traverse these organisations and are influenced by this counter-cultural, oppositional ethos – like the field of computer security or that of digital rights. For this reason, the incorporation of Big Tech in the state's security apparatus is marked with intense power struggles that are often made visible, for instance through the media.

Post-Snowden: cooperation or resistance?

These struggles can provide key insights to understand data politics and modern state power. In recent research conducted on the surveillance of Internet communications by intelligence agencies, we approached these issues by looking at the debates around Internet surveillance in the aftermath of the 2013 Snowden disclosures in the United States and in France (Tréguer 2018). By following interactions between Big Tech and governments as they moved from surveillance to other issues of interest to the security field (such as the weakening of encryption or the fight against terrorist propaganda), we worked through an inductive approach to identify factors influencing how these profit-seeking entities and their managers would fall in the cooperation/resistance spectrum, depending on the changing context and constraints that they face across time and space.

Among the factors making up this constraint structure were a firm's internal corporate culture, past and ongoing dealings with the human rights field (e.g. past human rights scandals affecting them), the importance of user trust and the threat

of competition. The relative weight of these constraints in a given context made resistance to the demands of the security field more likely. In turn, the sensitivity of these firms to regulatory changes, the identification of their managers to what Mills called the "power elite" (Mills 1959), their dependence on public funding and procurements, and the existence of criminal sanctions for non-cooperation all made cooperation more likely.

Looking at post-Snowden debates in the United States and in France to see how this constraint structure played out, we noticed some differences between the two countries from 2013 to 2015. In the US (and although these actions had global repercussions), we first see overt and multi-pronged resistance strategies being staged by Big Tech, whether through "technical resistance" with the roll-out of encryption on their products, or legal and political resistance through litigation and political advocacy aimed at reigning in the power of intelligence agencies.

In part, these can be read as instances of "double dealings" in the field of human rights defenders and that of hackers and engineers who were mobilised to beef up privacy protections in response to the Snowden disclosures. In Bourdieu's research such double-dealings refer to situations "whereby leaders, managers, officials or delegates of a field appear to be acting in a disinterested or principled manner 'for the field' and its values but are actually serving their own interests" (Webb, Schirato, and Danaher 2002). By aligning themselves with the privacy claims of their own workers concerned about their incursion in the military-industrial complex, the demands of human rights organisations and those of the field of computer security, these firms were able to remobilise workers, mitigate reputational risk and restore the trust of their users and customers concerned by the revelations, thereby securing or even reinforcing their market positions.

Encryption is a case in point. Whereas media coverage often over-emphasised the tensions between Big Tech and governments on this issue – for instance in 2016 when Apple refused a request by the FBI to collaborate in order to bypass encryption on a iPhone used by the San Bernadino shooter – such legal resistance often simply came down to respecting the state of the art, as computer engineers across the world worked to draw the lessons of the Snowden disclosures and beef up computer security globally (Rogers and Eden 2017).

Big Tech did not go much further. When strong, end-to-end encryption was rolled-out, the companies often declined to make it a by-default option. It was the case with Facebook Messenger, Microsoft's Skype or Google's Allo. An FBI source reacted to the launch of Allo by saying that "having [strong encryption] as an opt-in feature is certainly useful to us" (as quoted in Nakashima and Tsukayama 2016). Even when they are used by default, the strong encryption features like those deployed by Facebook on WhatsApp only encrypt the content of communications, not the metadata (who communicates with whom, when, from where, etc.). In other words, these deployments still allowed companies to mine metadata so as to monitor their users' behaviour and serve them with targeted advertising. Of course, such metadata can be, and frequently is, handed over to law enforcement (e.g. Biddle 2016; Fox-Brewster 2017).

Looking at these developments, some scholars have argued that the spread of encryption on Big Tech's infrastructures – which, according to NSA officials, had a significant inhibiting effect on the surveillance capabilities of law enforcement and intelligence agencies (McLaughlin 2016) – can be seen as a way of ensuring that the surveillance of users' communication could only happen with the companies' knowledge and consent, thereby reinforcing their position in the security field (Rubinstein and Van Hoboken 2014).

These shortcomings may primarily be driven by business considerations, rather than result from direct negotiations between Big Tech and the security field. But in mid-2015, a White House memo on encryption contemplated the possibility of "voluntary assistance," possibly in a "private" way to avoid the chilling effects that publicity might have on such cooperation (US National Security Council 2015). Since then, US intelligence and its allies have indeed exerted more quiet pressure to boost cooperation (e.g. Sanger and Frenkel 2018). We also know from the Snowden archives that prior to 2013, the NSA spent $250 million a year to work with tech companies to make commercial software – and in particular encryption software – more exploitable (Ball, Borger, and Greenwald 2013).

Surveillance reform and the Snowden paradox

When it comes to legal and political resistance staged by Big Tech, they too have their limits. The case of the US indeed confirms that despite an increased degree of transparency, surveillance reform introduced since 2013 in liberal regimes has led to what we have called the "Snowden paradox" (Tréguer 2017): Intelligence reform, rather than rolling-out capacities for large-scale and "suspicionless" surveillance, has provided a detailed legal basis for these capacities, bringing a few new safeguards and decreasing the level of secrecy to secure their legality and legitimacy.

In the US, the most important piece of legislation in this respect was the USA Freedom Act, passed in June 2015. Rather than allowing the NSA to collect and store domestic telephone records in bulk, the legislation effectively gives that authority to telecommunication providers (who will have to query their own databases with selectors provided by the NSA and hand over the matching data). In no way did this stop the growth of large-scale surveillance conducted by US intelligence. According to the reports published by the Office of the Director of National Intelligence in 2017 and 2018, the amount of data collected by the NSA has surged since 2013 (Gallagher 2017; Volz 2018), including the data collected from Big Tech (according to the Google Transparency Report, by virtue of the Foreign Intelligence Surveillance Act, it provided data on 14,000 user accounts in the first semester of 2013; in the first half of 2017, that number rose to more than 48,500 accounts – a 350% increase). Even legislative changes allowing companies to ask a judge to review gag orders attached to surveillance requests (preventing any public disclosure on the existence of such requests) have since been selectively used by companies like Google and Facebook, leading to criticisms from human rights organisations (Cardozo 2017).

In France, US tech companies' push for surveillance reform originally acted in a much more antagonist environment, partly resulting from existing public-private alliances with French telecom and defence firms and from repeated calls in favour of "digital sovereignty" by decreasing the dependency of the French security field on US companies. In fact, Big Tech's initial attempt at resistance to the expansion of state surveillance capacities was immediately denounced by government officials as hypocritical, considering their own commercial surveillance practices. This subsequently led to much less intense and more discreet forms of engagement when France expanded the surveillance powers of its intelligence agencies through this new legislation. As for the amount of data provided by Big Tech to the French police and judiciary in the past years, it has also sharply increased, in part due to a better compliance rate after a "group of contact" was established between technology companies and the Ministry of the Interior in 2015 (Cassini 2015). In the first half of 2013 (January–June), Google was served with 2,011 requests by French authorities (it complied with 49 % of them); Facebook was served with 1,547 requests (39 % compliance rate). In the first half of 2017, Google received 5,661 requests (it complied with 63 % of them); Facebook, 4,700 requests (74%). In four years, that makes for a 360% and 570% increase in the number of requests for which some data was produced, respectively.

Controlling data flows: A "fundamental shift" in "scale and nature"

After 2015 and save for a few exceptions, the influence of the human rights field on the global debate on surveillance reform withered along with media attention to state surveillance issues. Through securitisation discourses – where securitisation refers to speech acts calling for urgent and exceptional measures to deal with the terrorist threat (Buzan and Wæver 2003, 491) – the security agenda became dominated by the terrorist threat. This led to new calls on the part of the security field to limit encryption, boost surveillance capabilities, and fight against terrorist propaganda on online networks, often with the threat of new legislation and criminal sanctions if the companies failed to cooperate. In this context, similar trends towards greater cooperation materialised both in the US and Europe, suggesting strong transnational field effects across the Atlantic (Bigo 2016). Two topics are particularly illustrative.

One area of cooperation that will help increase the already impressive growth in data requests sent to Big Tech companies are ongoing reforms around extraterritorial access to data. In March 2017, US President Donald Trump signed into law the CLOUD Act. This piece of legislation – first presented by the Department of Justice in mid-2016 – was added at the last minute to a spending bill to revise the legal framework regulating US law enforcement access to online data stored overseas as well as access to data by foreign law enforcement authorities to data held in the US.

The goal is to provide a streamlined legal avenue to bypass the often long and tedious procedures of international judicial cooperation provided by Mutual Legal Assistance Treaties (MLATs) (Vergnolle 2017) while clarifying

the extra-territorial effects of US law. With the CLOUD Act, if a country is deemed by the US government to have an adequate legal framework for surveillance, a bilateral agreement will be concluded giving to that country the possibility to directly send surveillance requests to US companies, without having to go through the US judiciary, even when the data is stored in the US. Conversely, US law enforcement agencies will be able to request any user data from US companies, regardless of the nationality of targeted persons and regardless of where the data is stored.

Drafted and passed with wide-ranging support from the tech sector (Smith B, 2018; Walker 2017), the CLOUD Act further entrenches the privatisation of the justice system for regulating trans-border data flows. As a result of this legislation, tech companies who receive requests from a third country will be the only ones able to oppose these requests; with MLAT procedures, the whole process would have been supervised by foreign and US judges. It also gives new leverage to the US government – i.e. whether or not to conclude a bilateral agreement with foreign governments to give them direct access to the data troves of US companies – which may be abused to further the diplomatic interests of the US at the expense of human rights. These are just some of the most obvious problems of a legislation reaped with ambiguities (Singh Guiliani and Shah 2018; Wong 2017). In Europe, the adoption of the CLOUD Act was immediately followed by a proposal for a "directive on electronic evidence" aimed at enacting similar rules. Both initiatives could quickly expand worldwide through an ongoing revision of the Convention on Cybercrime of the Council of Europe.

In both the US and Europe, another hot topic has been that of terrorist propaganda, and the intensifying pressure put on online service providers to police such speech. Since the Paris attacks of 2015 and a visit of the French Minister of the Interior to Silicon Valley, France has been the European leader in this push toward privatised censorship, which was quickly taken up at the level of the European Union. The European Commission and Europol have convened regular meetings to get online platforms to sign a code on hate speech in 2016. In its report on the activity of its "Internet Referral Unit" created in 2015 to weed out extremist content online, Europol makes clear that these censorship activities are conducted outside of any legislative framework:

> A referral activity (meaning the reporting of terrorist and extremist online content to the concerned online service provider) does not constitute an enforceable act. Thus, the decision and removal of the referred terrorist and extremist online content is taken by the concerned service provider under their own responsibility and accountability (in reference to their Terms and Conditions).
>
> *(Europol 2016, 4)*

The US government followed suit in February 2016, when US Cabinet members and intelligence officials also met with tech companies. At the time, the White

House press secretary told reporters that "many of these technology companies that are participating in the meeting today are run by patriotic Americans and would want to cooperate" (as quoted in Jose and California 2016). A key aspect of the discussions laid in understanding how technology could be used to boost censorship of terrorist propaganda and so-called counter-discourse. A few months later, companies like Google and Facebook were announcing major innovations in their efforts on extra-judicial automated censorship, something that the government could not do considering the "First amendment" issues raised by such policies (Menn and Volz 2016).

Of course, online censorship represents an important challenge considering the sheer volume of third-party content posted on these platforms: 300 hours of video are posted on YouTube every minute; and on Facebook, every day, 510,000 comments are posted, 293,000 statuses are updated, and 136,000 photos are uploaded. So besides hiring thousands of content-moderators, often through subcontractors based in low-wage countries where basic social rights are discarded (Roberts 2016), these efforts have led to significant investment in tools based on "Machine Learning" systems aimed at censoring terrorist-related content.

As the British Prime Minister Theresa May, explained at UN General Assembly in New York in September 2017:

> Industry needs to go further and faster in automating the detection and removal of terrorist content online, and developing technological solutions which prevent it being uploaded in the first place. We need a fundamental shift in the scale and nature of our response – both from industry and governments.
>
> *(as quoted in Hope and McCann 2017)*

A year later, in September 2018, the European Commission was announcing a proposal for transcribing the extra-judicial and automated mechanisms experimented over the past years into EU law.

What is happening here, with these moves around extra-territorial access to data or the censorship of terrorist propaganda, is a major reconfiguration of security assemblages tasked with the management of data flows and stocks, as new actors, technologies and regulations become necessary for the state to handle the surge in public and private communications entailed by digital technologies and keep its traditional techniques of power afloat. But the more we look at these new security assemblages involving Big Tech, the more we start to understand that they are but a sign of a broader reconfiguration of bureaucracies in the digital age.

Bureaucracies in the age of data governance

In her work on neo-liberal bureaucratisation, Béatrice Hibou has shown how, from the 1970s on, the logic of management migrated from the private realm to state

institutions (Hibou 2015). Through imperatives of "efficiency", "cost-effectiveness", "flexibility" and through practices of "auditing", and "benchmarking", bureaucratic practices within public administrations grew increasingly hostile to the post-war social values embedded in the public sector. With the so-called "New Public Management," the abstract principles of neoliberalism were pushed so far as to cause a complete divorce between "efficiency" and the ends that state bureaucracies were supposed to pursue (Graeber 2015).

If, according Hibou, bureaucracy is seen "as a power concentrated in the hands of those who create the validated abstractions and put them at the core of government" (p. 86), it looks like Big Tech may fast be dominating the whole administrative field. Zuboff (2015) and others trace back the origin of this diffusion of power to Google's popularisation of "data governance", referring to bureaucratic models based on "data extraction and analysis", "new contractual forms due to better monitoring", "personalization and customization", and "continuous experiments" championed by Google's chief economist (Varian 2010, 2014). They were later relayed in the book *How Google Works?* (2015), authored by Eric Schmidt, Executive Chairman of Google/Alphabet from 2001 to 2017 and still a board member today, and Jonathan Rosenberg, a former Senior Vice President of Products at Google. In this book, the pair document business management lessons from Google, an experience that led them to "relearn everything" they knew through data-intensive models.

Data means knowledge means hard power

After having contributed to the shaping of US Internet diplomacy under the tenure of former Secretary of State Hillary Clinton (Assange 2014; Powers and Jablonski 2015; Schmidt and Cohen 2013), Eric Schmidt has been the most visible agent of a crowd of current and former "Googlers" helping spread these models at the heart of the US military-industrial complex. In March 2016, Schmidt was appointed by the Secretary of Defense as chairman of the Defense Innovation Board (DIB). A position as advisor to the Pentagon that he still holds at the time of writing despite him quitting his official positions at Alphabet/Google in January 2018. On its webpage, the DIB is described as an innovation think-tank:

> Through pilot programs and experiments within DoD, the DIB can bring in new perspectives from the private sector and academia, work with DoD partners to test hypotheses, gather data, and encourage the imagination and critical thinking need to consider new solutions. This process is rapid, creative, collaborative, and ultimately saves time and money.
>
> *(DIB 2017)*

In one of its recommendations entitled "Forge New Approach to Data Collection, Sharing, and Analysis", the DIB insists on the importance of data to 21st century statehood:

> Data is the 21st century equivalent of a global natural resource, like timber, iron, or oil previously – indispensable for sustaining military innovation and advantage. The next global conflicts will be fueled by data. The rapidly expanding power of new mathematical and computing techniques to reveal insights into intentions and capabilities, and to enhance accuracy, lethality, and speed, depend on immense data sets to train algorithms and from which to extract information. The data that provide the raw materials from which to identify patterns, as well as the anomalies that defy them, constitute the fuel that powers the engine of Machine Learning (ML). Whoever amasses and organises the most data first will sustain technological superiority, so it is incumbent upon the Department to collect, store, share, analyze, and protect its data faster and better than its competitors. Data must be regarded as one of the most powerful resources in the Department's arsenal.
>
> *(DIB 2017)*

Companies like Google and their executives are selling solutions aimed at expanding the technological superiority and "efficiency" of security bureaucracies. According Scott Frohman, Google's Director of Defense and Intelligence Sales, Big Tech can bring these "radical innovations" at "ultra-low cost." "Through the use of Google's capabilities remade for the enterprise," he writes on his LinkedIn profile, "the government gets innovation fast and with significantly reduced cost" (Frohman 2018).

The "Startup Nation" as a new bureaucratic paradigm

These trends go beyond the security field and expand to virtually all public policies. The integration of Big Tech in the administrative and political fields has been going on for at least a decade. It vastly expanded under the Obama administration, with over 251 individuals changing position between Google or related firms and the federal government, national political campaigns and Congress ('Google's Revolving Door (US)' 2017).

Under Trump, it may look different on the surface. For one thing, the tech industry has voiced strong criticism of his immigration and climate policies (Streitfeld, Isaac and Benner 2017). Big Tech workers have also played an important role in denouncing Google's participation to drone warfare. At Microsoft they opposed a $19.4 million contract with US Immigration and Customs Enforcement (ICE), while Amazon was criticised for selling facial recognition technologies to US police forces. Trump has of course "trolled" Big Tech, for instance by accusing them of censoring conservative views online (Swisher 2018).

But in the back rooms, it looks like business as usual. In a memorandum signed in late-March 2017 creating the "American Technology Council," Donald Trump opened new channels for sustaining the reciprocal influence between the tech industry and the US government. The initiative is overseen by his son-in-law and Senior Advisor Jared Kushner and seeks to "modernize" the US public sector. Discussions have touched on how to make public procurement more flexible, cut

down on some 6,000 government-owned data centres by shifting those responsibilities to the private sector, or on the release government-held data on a range of issues, particularly on health care, for private-sector use (Romm 2017).

In France too, where a lot of revolving door activity is also happening ("Google's European Revolving Door" 2016; Léchenet 2017), the debate on the "reform of the state" has moved from the premises of the New Public Management to those of data governance. By coining terms like "Startup Nation" – an expression championed by French President Emmanuel Macron – or the "Platform State," today's reformers are re-modelling bureaucracies and decision-making processes around the need to produce massive amounts of data, make it available and usable, maintain its integrity and feed it to powerful data-processing tools that will be used to "optimise" bureaucratic outputs (e.g. Algan and Cazenave 2016; Bertholet and Létourneau 2017; Pezziardi and Verdier 2017). Even when these reformist discourses claim to be opposing the hegemony of US tech companies, they are in fact assuming the superiority of their models and diffusing them across public administrations. It is an instance of "mimetic rivalry" (Girard 2002), where what Evgeny Morozov has termed solutionism serves as a new technocratic utopia (Morozov 2013).

Despite calls of security insiders and state reformers to establish "digital sovereignty," the products and services of US tech firms continue to have an irresistible appeal. In 2017, a contract between Microsoft and the French Ministry of Defence was signed despite widespread criticism. A year earlier, the DGSI, France's domestic intelligence, contracted Palantir, a Big Data analytics firm very close to US intelligence, to mine the vast amount of data seized during house raids and digital seizures authorised under the state of emergency post-November 2015 (Tesquet 2017).

As this latter example suggests, new security assemblages do not only embark large tech firms used daily by billions of Internet users. Many small companies specialised in data analytics or vulnerabilities are also partnering with intelligence agencies to sell their products and services (Deibert 2013). Older tech and utility companies in the defence, transportation or energy sectors are also trying to catch up by investing in Big Data analytics – sometimes in partnership with their US competitors like IBM to secure access to key technologies – and are fast-developing solutions for Big Data policing, just as local elected officials hope to get votes by framing these new programs as advancing the project of a "Smart City."

The Government Machine and the rule of law

Such trends towards technical, managerial and technological responses to security challenges have been sweeping the modern security field for quite some time now (Abrahamsen and Williams 2010; Bonelli 2010). But as Big Tech becomes part of the state and now serves the "Government Machine" (Agar 2003), we might be reaching a tipping point in the history of governmentality.

That being said, political theory suggests that the blurring public-private distinction is a feature of state power, not a bug. According to Timothy Mitchell, we need to see the state not "as a free-standing entity, whether an agent, instrument,

organisation or structure, located apart from and opposed to another entity called society," but rather as a multiplicity of political arrangements that produce structural effects that maintain social and political order (Mitchell 1991, 94). From this perspective, "the boundary of the state is merely the effect of such arrangements and does not mark a real edge. It is not the border of an actual object." Rather, "producing and maintaining the distinction between state and society is itself a mechanism that generates resources of power."

The "Big Tech vs. the Surveillance State" narrative emphasises that distinction. Post-2013, it served to counter that put forward by Snowden and journalists working on his disclosures of unabated and extra-legal cooperation. It helped reassure Internet users that these companies worked to protect their rights and resisted the state on their behalf. But soon enough, through more discreet moves, the state-private distinction was again crushed when, to effectively control communications and avoid investing resources in the justice system, security professionals co-opted Big Tech and their censorship and surveillance techniques for their own ends. Such double-dealings are still ongoing.

Towards hybrid rule

Through public-private hybridisation, it becomes easier for governments to escape the important safeguards that our legal systems have developed over time to protect political rights, but which are apparently ill-suited (or at least too costly) to accommodate the surge in communications entailed by digital technologies. In this way, "political elites (. . .) rely on the private sector to shield national security activities," thus "expanding state power while constraining democratic accountability" (Hurt and Lipschutz 2015, 2).

This was in part a deliberate strategy first envisioned in the mid-1990s, when security professionals – in particular at the Pentagon – feared the consequence of the compression of time and space induced by digital networks and sought for new ways of enacting state power. The result was the formulation of new doctrines and practices whereby "the military and law enforcement, the government and private industry, and domestic and foreign surveillance would necessarily mix in ways long seen as illicit if not illegal" (Jones 2017, 13). It followed that "constitutional interpretation, jurisdictional divisions, and the organisation of bureaucracies alike would need to undergo dramatic – and painful – change."

As we have seen, part of such change resides in automation and extra-judicialisation,which both lead to a profound shift in the history of the justice system. In his 1971–1972 lectures at the Collège de France, Foucault explained how, from the 14th century on, the old feudal institution of Parliaments was gradually co-opted by the Crown and entrusted with investigative powers (Foucault 2018). To assert its power without having to bear the costs of military occupation, the Prince relied on Parliaments to interrogate people, to make them say what they knew so as to produce knowledge on the basis of which he would adjudicate and eventually govern "his" territory and the population.

In today's computerised world, the legal restrictions of state power that were progressively coded into the justice system are radically overtaken by "data governance." Digital traces form the basis of a new statistical power-knowledge that is seen as the most effective and cost-efficient way of governing the natural and social worlds. In the process, the legal norms and principles which somewhat circumscribed the power of the Prince get lost in computer code. Once it has morphed into algorithms, power becomes even more diluted and harder to challenge (Rouvroy, 2012). Can we even reasonably hope to make these ever more complex algorithms auditable, and their designers accountable, when experts in "Deep Learning" and "Neural Networks" say that even they cannot understand how these increasingly unpredictable systems work (Knight 2017; Smith, A. 2018)?

Stopping the machine?

That begs the question of how best to challenge these new security assemblages.

For one thing, it is worth stressing that Big Data bureaucracies might not be that good at doing what they are supposed to. We can therefore oppose the arguments of those legitimising these new governance models on the grounds of accuracy and reliability. There are reasons – and growing evidence – to doubt that the new "regimes of truth" championed by "Big Data security assemblages" will in any meaningful way provide solutions to security issues (Aradau and Blanke 2015). Technological solutionism in the age of data governance, bolstered by marketing discourses, might only be recreating a veil of illusion of technocratic control, while putting evermore distance between bureaucracies and the social world they wish to make more orderly. After all, history tells us, bureaucracies tend to fail. By disregarding "all the subtleties of real social existence," "reducing everything to preconceived mechanical or statistical formulae," bureaucratic dispositifs like "forms, rules, statistics, or questionnaire" – even when fuelled by complex algorithms and troves of data –, remain abstract simplifications that might only reinforce the forms of structural violence they are said to alleviate or even solve (Graeber 2015, p. 75; see also Eubanks 2018, O'Neil 2016).

Bureaucracies often fail to meet their alleged goals but still, they strengthen the power of those who invest in them. They transform the social world and can go awfully wrong (Scott 1998). If, following Tim Mitchell and critical security scholars, we refuse to "see the state and private organizations as a single, totalized structure of power," another complementary way of resisting these assemblages is to build on the conflicts that inevitably occur "between different government agencies, between corporate organizations, and within each of them" (Mitchell 1991, 90). We can amplify the words of those who denounce the oppressive and manipulative use of modern computer technologies, applaud tech workers opposing the direct involvement of their company in the military-industrial complex, or support security professionals seeking to automate intelligence oversight so as to catch up with large-scale surveillance systems and mitigate abuse. We can, and we should.

But post-Snowden controversies also show that these forms of resistance create a risk that we will overlook the pervasiveness of the institutions and technologies, of the rationalities and practices that created the problem in the first place. The risk is that all we are able to come up with are legal, technological or bureaucratic fixes to try to contain the most disturbing aspects of data-driven bureaucracies, without affecting the longer-term trend of a technological arms race that only seems to intensify the issues it was allegedly meant to solve.

Some kind of deeper resistance might be warranted. In his writings on power, Foucault once asked: "How can the growth of capabilities" – and he explicitly mentioned "techniques of communication – "be disconnected from the intensification of power relations?" (Foucault 1984, 48). Computing technologies have since become immensely powerful and yet, we are still struggling to find a satisfactory answer to this crucial question. Despite the hopes of early hackers and Internet pioneers, the decoupling of technology and power is not happening. The key question then becomes whether technology itself or law or ethics can actually be effective instruments to achieve such decoupling.

At this stage of technological development, if we feel like they cannot – at least not in the near foreseeable future – that means it is probably time to refocus on tackling the imaginaries and institutions that underlie the "growth of capabilities" itself: the blind faith in technological progress; the oft-repeated mantra that technology is neutral, that its negative potential will somehow be contained; places like the universities, R&D labs, ministries, start-ups, shops and factories where complex and powerful technologies are designed, produced and traded. As philosopher Jacques Ellul observed, "we set huge machines in motion in order to arrive nowhere" (Ellul 1989, 51). If it is not Thomas More's eu-topia ("no place") that we are fast approaching but rather a "dystopian void" (The Luddbrarian 2018), what we need is not just a technological fix, a bureaucratic patch, a principled law or even a good ethics; what we need first and foremost is to get off and stop the machine.

References

Abrahamsen, Rita, and Anna Leander. 2015. *Routledge Handbook of Private Security Studies*. Routledge.

Abrahamsen, Rita, and Michael C. Williams. 2010. *Security Beyond the State: Private Security in International Politics*. Cambridge University Press.

Agar, Jon. 2003. *The Government Machine: A Revolutionary History of the Computer*. MIT Press.

Algan, Yann, and Thomas Cazenave. 2016. *L'Etat en mode start-up*. Eyrolles.

Aradau, Claudia, and Tobias Blanke. 2015. 'The (Big) Data-Security Assemblage: Knowledge and Critique'. *Big Data & Society* 2 (2). 1–12.

Assange, Julian. 2014. *When Google Met Wikileaks*. OR Books. https://www.orbooks.com/catalog/when-google-met-wikileaks/.

Ball, James, Julian Borger, and Glenn Greenwald. 2013. 'Revealed: How US and UK Spy Agencies Defeat Internet Privacy and Security'. *The Guardian*, 6 September 2013, sec. World news. http://www.theguardian.com/world/2013/sep/05/nsa-gchq-encryption-codes-security.

Barbrook, Richard, and Andy Cameron. 1995. 'The Californian Ideology'. *Mute* 1 (1): 44–72.

Barty-King, Hugh. 1980. *Girdle Round the Earth: History of Cable and Wireless*. Heinemann.

Bellamy Foster, John, and Robert W. McChesney. 2014. 'Surveillance Capitalism: Monopoly-Finance Capital, the Military-Industrial Complex, and the Digital Age'. *Monthly Review* 66 (3). http://monthlyreview.org/2014/07/01/surveillance-capitalism.

Beniger, James R. 1986. *The Control Revolution: Technological and Economic Origins of the Information Society*. Harvard University Press.

Bertholet, Clément, and Laura Létourneau. 2017. *Ubérisons l'État! Avant que d'autres ne s'en chargent*. Armand Colin.

Biddle, Sam. 2016. 'Apple Logs Your IMessage Contacts—and May Share Them with Police'. *The Intercept*. 28 September 2016. https://theintercept.com/2016/09/28/apple-logs-your-imessage-contacts-and-may-share-them-with-police/.

Bigo, Didier. 2016. 'Sociology of Transnational Guilds'. *International Political Sociology* 10 (4): 398–416.

Black, Edwin. 2012. *IBM and the Holocaust: The Strategic Alliance Between Nazi Germany and America's Most Powerful Corporation*. Expanded edition. Washington, DC: Dialog Press.

Bonelli, Laurent. 2010. 'Les modernisations contradictoires de la police nationale'. In *L'État démantelé*, 102–17. La Découverte.

Buzan, Barry, and Ole Wæver. 2003. *Regions and Powers: The Structure of International Security*. Cambridge University Press.

Cardozo, Nate. 2017. 'Requiring Judicial Review for Every Gag Order Is a Simple Way to Have Our Backs: Apple Does but Google and Facebook Fall Short'. *Electronic Frontier Foundation*. 10 July 2017. www.eff.org/deeplinks/2017/07/requiring-judicial-review-every-gag-order-simple-way-have-our-backs-apple-does.

Cassini, Sandrine. 2015. 'Terrorisme: Accord Entre La France et Les Géants Du Net'. *Les Échos*. 23 April 2015. www.lesechos.fr/journal20150423/lec2_high_tech_et_medias/02124922454-terrorisme-accord-entre-la-france-et-les-geants-du-net-1113723.php.

Charbon, Paul. 1991. 'Genèse du vote de la loi de 1837, origine du monopole des télécommunications'. In *L'État et les télécommunications en France et à l'étranger*, 18371987, edited by Catherine Bertho-Lavenir, 11–22. Actes du colloque organisé à Paris les 3 et 4 novembre 1987 par l'Ecole pratique des hautes études et l'Université René Descartes. Librairie Droz.

Deibert, Ronald J. 2013. *Black Code: Inside the Battle for Cyberspace*. Random House.

Desrosières, Alain. 2002. *The Politics of Large Numbers: A History of Statistical Reasoning*. Translated by Camille Naish. Harvard University Press.

DIB. 2017. 'Recommendations'. *Defense Innovation Board*. 2017. https://innovation.defense.gov/Recommendations/.

Edwards, Paul N. 1996. *The Closed World: Computers and the Politics of Discourse in Cold War America*. MIT Press.

Ellul, Jacques. 1989. The Presence of the Kingdom. 2nd edition. Helmers & Howard Publishing.

Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.

Europol. 2016. 'EU Internet Referral Unit – YEAR ONE REPORT'. *Europol*. www.europol.europa.eu/content/eu-internet-referral-unit-year-one-report-highlights.

Flichy, Patrice. 2009. *Dynamics of Modern Communication: The Shaping and Impact Of New Communication Technologies*. SAGE.

Foucault, Michel. 1984. 'What Is Enlightenment?' In *The Foucault Reader*, edited by Paul Rabinow. Pantheon Books.

———. 2018. *Penal Theories and Institutions: Lectures at the Collège de France*, 1971–1972. Translated by Graham Burchell. 1st edition. 2018 edition. Palgrave Macmillan.

Fox-Brewster, Thomas. 2017. 'Forget About Backdoors, This Is the Data WhatsApp Actually Hands to Cops'. *Forbes*. 22 January 2017. www.forbes.com/sites/thomasbrewster/2017/01/22/whatsapp-facebook-backdoor-government-data-request/.

Frohman, Scott. 2018. 'Scott Frohman LinkedIn Profile'. *LinkedIn*. 2018. https://www.linkedin.com/in/scottatsas/.

Fuchs, Christian, and Daniel Trottier. 2015. 'Towards a Theoretical Model of Social Media Surveillance in Contemporary Society'. *Communications-European Journal of Communication Research* 40 (1): 113–35.

Gallagher, Sean. 2017. 'US Intelligence "Transparency Report" Reveals Breadth of Surveillance by NSA, Others'. *Ars Technica*. 3 May 2017. https://arstechnica.com/tech-policy/2017/05/us-intelligence-transparency-report-reveals-breadth-of-surveillance-by-nsa-others/.

Galloway, Alexander R. 2004. *Protocol: How Control Exists after Decentralization*. MIT Press.

Girard, René. 2002. 'What Is Happening Today Is Mimetic Rivalry on a Global Scale'. *South Central Review* 19 (2/3): 22–27.

'Global Top 100 Companies by Market Capitalisation (31 March 2017 Update)'. 2017. https://www.pwc.com/gx/en/audit-services/assets/pdf/global-top-100-companies-2017-final.pdf.

Goldsmith, Jack, and Tim Wu. 2006. *Who Controls the Internet?: Illusions of a Borderless World*. Oxford University Press.

'Google's European Revolving Door'. 2016. *Google Transparency Project*. 2016. https://googletransparencyproject.org/articles/googles-european-revolving-door.

'Google's Revolving Door (US)'. 2017. *Google Transparency Project*. 2017. https://googletransparencyproject.org/articles/googles-revolving-door-us.

Graeber, David. 2015. *The Utopia of Rules: On Technology, Stupidity, and the Secret Joys of Bureaucracy*. Melville House.

Haggerty, Kevin D., and Richard V. Ericson. 2000. 'The Surveillant Assemblage'. *British Journal of Sociology* 51 (4): 605–22.

Harcourt, Bernard E. 2015. *Exposed - Desire and Disobedience in the Digital Age*. Harvard University Press.

Harris, Shane. 2014. *@War: The Rise of the Military-Internet Complex*. Eamon Dolan/Houghton Mifflin Harcourt.

Headrick, Daniel R. 2012. *The Invisible Weapon: Telecommunications and International Politics, 1851–1945*. Reprint. Oxford University Press, USA.

Hibou, Béatrice. 2015. *The Bureaucratization of the World in the Neoliberal Era*. Palgrave MacMillan.

Hope, Christopher, and Kate McCann. 2017. 'Google, Facebook and Twitter Told to Take down Terror Content within Two Hours or Face Fines'. *The Telegraph*, 19 September 2017. https://www.telegraph.co.uk/news/2017/09/19/google-facebook-twitter-told-take-terror-content-within-two/.

Hurt, Shelley, and Ronnie Lipschutz, eds. 2015. *Hybrid Rule and State Formation: Public-Private Power in the 21st Century*. 1st edition. Routledge.

Jones, Matthew L. 2017. 'The Spy Who Pwned Me'. *Limn*, no. 8 (June). http://limn.it/the-spy-who-pwned-me/.

Jose, Danny Yadron Julia Carrie Wong in San, and California. 2016. 'Silicon Valley Appears Open to Helping US Spy Agencies after Terrorism Summit'. *The Guardian*, 8 January 2016, sec. Technology. http://www.theguardian.com/technology/2016/jan/08/technology-executives-white-house-isis-terrorism-meeting-silicon-valley-facebook-apple-twitter-microsoft.

Knight, Will. 2017. 'There's a Big Problem with AI: Even Its Creators Can't Explain How It Works'. *MIT Technology Review*. 11 April 2017. https://github.com/alphoenix/donnees/tree/master/lobbies-gafamut.

Léchenet, Alexandre. 2017. 'L'influence Tentaculaire Des Géants Américains'. https://github.com/alphoenix/donnees.

Lécuyer, Christophe. 2007. *Making Silicon Valley: Innovation and the Growth of High Tech, 1930–1970*. MIT Press.

Levine, Yasha. 2018. *Surveillance Valley: The Secret Military History of the Internet*. PublicAffairs.

Luddbrarian, The. 2018. 'Challenging the Tech Companies from Within'. LibrarianShipwreck, 28 June 2018. https://librarianshipwreck.wordpress.com/2018/06/28/challenging-the-tech-companies-from-within/.

McLaughlin, Jenna. 2016. 'Spy Chief Complains That Edward Snowden Sped Up Spread of Encryption by 7 Years'. *The Intercept*. 25 April 2016. https://theintercept.com/2016/04/25/spy-chief-complains-that-edward-snowden-sped-up-spread-of-encryption-by-7-years/.

Menn, Joseph, and Dustin Volz. 2016. 'Exclusive: Google, Facebook Quietly Move toward Automatic Blocking of Extremist Videos'. *Reuters*, 25 June 2016. www.reuters.com/article/us-internet-extremism-video-exclusive-idUSKCN0ZB00M.

Mills, C. Wright. 1959. *The Power Elite*. Oxford University Press.

Mitchell, Timothy. 1991. 'The Limits of the State: Beyond Statist Approaches and Their Critics'. *The American Political Science Review* 85 (1): 77.

Morozov, Evgeny. 2013. *To Save Everything, Click Here: The Folly of Technological Solutionism*. PublicAffairs,U.S.

Nakashima, Ellen, and Hayley Tsukayama. 2016. 'Why People like Edward Snowden Say They Will Boycott Google's Newest Messaging App'. *Washington Post*. 21 May 2016. www.washingtonpost.com/news/the-switch/wp/2016/05/21/why-people-like-edward-snowden-say-they-will-boycott-googles-newest-messaging-app/.

Nesbit, Jeff. 2017. 'Google's True Origin Partly Lies in CIA and NSA Research Grants for Mass Surveillance'. *Quartz*. 8 December 2017. https://qz.com/1145669/googles-true-origin-partly-lies-in-cia-and-nsa-research-grants-for-mass-surveillance/.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

Pezziardi, Pierre, and Henri Verdier. 2017. *Des startups d'État à l'État plateforme*. CreateSpace Independent Publishing Platform.

Powers, Shawn M., and Michael Jablonski. 2015. *The Real Cyber War: The Political Economy of Internet Freedom*. 1st edition. University of Illinois Press.

Roberts, Sarah T. 2016. 'Commercial Content Moderation: Digital Laborers' Dirty Work'. In *The Intersectional Internet: Race, Sex, Class, and Culture Online*, edited by Brendesha M. Tynes and Safiya Umoja Noble, 147–60. Peter Lang Publishing.

Rogers, Michael, and Grace Eden. 2017. 'The Snowden Disclosures, Technical Standards, and the Making of Surveillance Infrastructures'. *International Journal of Communication* 11 (0): 22.

Romm, Tony. 2017. 'Behind the Scenes at President Trump's Private Talks with the Tech Industry'. *Recode*. 20 June 2017. www.recode.net/2017/6/20/15838646/trump-apple-amazon-google-microsoft-tech-week.

Rouvroy, Antoinette. 2012. 'The End(s) of Critique: Data-Behaviourism vs. Due-Process'. In *Privacy, Due Process and the Computational Turn*. Routledge. http://works.bepress.com/antoinette_rouvroy/44.

Rubinstein, Ira, and Joris Van Hoboken. 2014. 'Privacy and Security in the Cloud: Some Realism About Technical Solutions to Transnational Surveillance in the Post-Snowden Era'. *Maine Law Review* 66 (2): 488–533.

Sanger, David E., and Sheera Frenkel. 2018. '"Five Eyes" Nations Quietly Demand Government Access to Encrypted Data'. *The New York Times*. 5 September 2018. www.nytimes.com/2018/09/04/us/politics/government-access-encrypted-data.html.

Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press.

Schmidt, Eric, and Jared Cohen. 2013. *The New Digital Age: Transforming Nations, Businesses, and Our Lives*. Knopf Doubleday Publishing Group.

Schmidt, Eric, and Jonathan Rosenberg. 2015. *How Google Works*. London: John Murray.

Singh Guiliani, Neema, and Naureen Shah. 2018. 'The CLOUD Act Doesn't Help Privacy and Human Rights: It Hurts Them'. *Lawfare*. 16 March 2018. https://lawfareblog.com/cloud-act-doesnt-help-privacy-and-human-rights-it-hurts-them.

Smith, Andrew. 2018. 'Franken-Algorithms: The Deadly Consequences of Unpredictable Code', *The Guardian*, 30 August 2018. www.theguardian.com/technology/2018/aug/29/coding-algorithms-frankenalgos-program-danger.

Smith, Brad. 2018. 'The CLOUD Act Is an Important Step Forward, but Now More Steps Need to Follow'. *Microsoft on the Issues* (blog). 3 April 2018. https://blogs.microsoft.com/on-the-issues/2018/04/03/the-cloud-act-is-an-important-step-forward-but-now-more-steps-need-to-follow/.

Statista. 2018. 'Most Valuable Companies in the World 2018'. May 2018. www.statista.com/statistics/263264/top-companies-in-the-world-by-market-value/.

Streitfeld, David, Mike Isaac, and Katie Benner. 2017. 'Silicon Valley's Ambivalence Toward Trump Turns to Anger'. *The New York Times*, 29 January 2017. https://www.nytimes.com/2017/01/29/technology/silicon-valleys-ambivalence-toward-trump-turns-to-anger.html.

Swisher, Kara. 2018. 'Trump's Ludicrous Attack on Big Tech'. *The New York Times*. 30 August 2018. www.nytimes.com/2018/08/29/opinion/trump-bias-google-twitter.html.

Tesquet, Olivier. 2017. 'Palantir, l'encombrant Ami Américain Du Renseignement Français'. 27 January 2017. www.telerama.fr/medias/palantir-big-data-renseignement,153229.php.

Tréguer, Félix. 2017. 'Intelligence Reform and the Snowden Paradox: The Case of France'. *Media and Communication* 5 (1): 17–28.

Tréguer, Félix. 2018. 'US Technology Companies and State Surveillance in the Post-Snowden Context: Between Cooperation and Resistance'. *UTIC Deliverable* 5. Paris: CERI. https://halshs.archives-ouvertes.fr/halshs-01865140.

Turner, Fred. 2006. *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*. University of Chicago Press.

US National Security Council. 2015. 'Draft Options Paper on Strategic Approaches to Encryption'. Washington DC. http://apps.washingtonpost.com/g/documents/national/read-the-nsc-draft-options-paper-on-strategic-approaches-to-encryption/1742/.

Varian, Hal R. 2010. 'Computer Mediated Transactions'. *American Economic Review* 100 (2): 1–10.

Varian, Hal R . 2014. 'Beyond Big Data'. *Business Economics* 49 (1): 27–31.

Vergnolle, Suzanne. 2017. 'Understanding the French Criminal Justice System as a Tool for Reforming International Legal Cooperation and Cross-Border Data Requests'. *Georgia Tech Scheller College of Business Research Paper*. Data Protection, Privacy and European Regulation in the Digital Age. https://papers.ssrn.com/abstract=2921364.

Volz, Dustin. 2018. 'Spy Agency NSA Triples Collection of U.S. Phone Records: Official . . .'. Reuters, 8 May 2018. www.reuters.com/article/us-usa-cyber-surveillance/spy-agency-nsa-collected-500-million-u-s-call-records-in-2017-a-sharp-rise-official-report-idUSKBN1I52FR.

Walker, Kent. 2017. 'Digital Security and Due Process: A New Legal Framework for the Cloud Era'. *Google*. 22 June 2017. www.blog.google:443/topics/public-policy/digital-security-and-due-process-new-legal-framework-cloud-era/.

Webb, Jen, Tony Schirato, and Geoff Danaher. 2002. *Understanding Bourdieu*. SAGE.

Williams, Michael C. 2010. 'The Public, the Private and the Evolution of Security Studies'. *Security Dialogue* 41 (6): 623–30.

Wong, Cynthia M. 2017. 'US Cross-Border Data Deal Could Open Surveillance Floodgates'. *Human Rights Watch*. 18 September 2017. www.hrw.org/news/2017/09/18/us-cross-border-data-deal-could-open-surveillance-floodgates.

Zuboff, Shoshana. 2015. 'Big Other: Surveillance Capitalism and The Prospects of an Information Civilization'. *Journal of Information Technology* 30 (1): 75–89.

PART III
Subjects

9

TOWARDS DATA JUSTICE

Bridging anti-surveillance and social justice activism[1]

Lina Dencik, Arne Hintz and Jonathan Cable

The Snowden leaks provided unprecedented insights into the operations of state-corporate surveillance and highlighted the indiscriminate nature of large-scale data collection across communication networks and platforms in Western democracies, most notably the US National Security Agency (NSA) and the British Government Communications Headquarters (GCHQ). The documents illustrated the intricate ways in which everyday communication is integrated into an extensive regime of surveillance that relies considerably on the "data-fication" of many aspects of social life. Ordinary users' social activities are "sucked up as data, quantified and classified, making possible real-time tracking and monitoring" (Lyon 2014, 4). This information infrastructure characterizes a particular mode of governance, one that is rooted in a political economy in which the prevailing logic is to predict and modify human behaviour as a means to produce revenue and market control; what Zuboff (2015) has described as "surveillance capitalism". Such data-driven forms of social organization have significant implications for citizenship (cf., Hintz, Dencik and Wahl-Jorgensen 2018; Isin and Ruppert 2015) and particularly for how citizens might intervene, challenge and resist this form of governance. As part of the interplay between data and politics, resistance can take several forms. Much onus has been on the collection and use of data, and prominence has been placed on the use of counter-surveillance technologies such as encryption or anonymisation tools along with a focus on advocacy pertaining to privacy and data protection amongst digital rights groups. This has provided windows of opportunity for technological developments and legislative changes that speak particularly to concerns with the implications of surveillance programmes for secure communication infrastructures and individual privacy (Hintz and Brown 2017; Rogers and Eden 2017). However, the degree to which such strategies and concerns have expanded towards the broader range of politically-active and interested

publics is less clear. Moreover, a common agenda around an engagement with the politics of or in data in which data is seen as "generative of new forms of power relations and politics at different and interconnected scales" (Ruppert, Isin and Bigo 2017, 2) is difficult to identify.

In this chapter we explore the relationship between these broader concerns and data by analysing responses to the Snowden leaks amongst political activists in the UK.[2] The chapter draws from a series of in-depth interviews with UK-based activists engaged in a range of social justice concerns, exploring attitudes and practices in relation to mass data collection and digital surveillance. Based on this research, we develop the concept of "data justice" as a way of reframing prominent understandings of data politics.

The chapter starts by outlining the implications of the Snowden leaks for political activists before discussing how resistance to data-driven surveillance has predominantly emerged in their aftermath. We argue that resistance in the datafied society post-Snowden has tended to focus on techno-legal responses relating to the development and use of encryption and policy advocacy around privacy and data protection. This presents a particular way of framing and engaging with data politics. In light of this, we examine how these types of practices are negotiated amongst political activists and outline the extent to which the activists we interviewed view such resistance as part of their social justice agendas. We observe a significant level of ambiguity around technological resistance strategies, while policy responses to the Snowden leaks have largely been confined within particular expert communities. In the final part of the chapter, we therefore propose a (re)conceptualization of resistance to data collection and use that can address the implications of this data-driven form of governance in relation to broader social justice agendas. To that end, we introduce the notion of data justice which, we argue, would help contextualize datafication, connect it to social and economic justice concerns, and thereby contribute to transforming the role of data politics in current civil society practice and, potentially, public debate. This is particularly significant in light of the central role of data-driven processes in contemporary capitalism.

The Snowden leaks and political activism

The revelations of programmes designed to "bulk" collect data on citizen engagement with digital infrastructures[3] indicate the extent to which contemporary forms of governance are increasingly based on the ability to monitor, track and potentially predict the behaviour of entire populations. This is part of a broader emphasis on the role of "big data" in current societies (Kitchin 2014) that highlights the surveillance implications of the "big data" discourse. As Lyon (2015) has argued, surveillance culture came prominently into view simultaneously with the intensified security-discourse following 9/11 and the so-called war on terror. In particular, the uncertainty of the form and nature of potential threats in such a political climate provides an apparent necessity and justification for limitless measures to be taken to

ward off any such possible dangers. The focus, therefore, moves to the operationalization of how to perceive of these potential threats, in which the apparatuses of surveillance play an integral role (Massumi 2015). In such circumstance, the rise of "surveillance society" marks a social context characterized by an increasing amount of surveillance taking place alongside an explosion in the possible methods and means for observing and monitoring people's behaviour (Lyon 2001).

A central concern for Snowden and others has been the extent to which extensive forms of monitoring lead to a "chilling effect" in society that stifles the possibilities for challenging institutions of power and advocating for social change. Although the theory of "chilling effects" has historically been difficult to empirically prove and remains controversial, the debate on it following the Snowden leaks concerned the extent to which government surveillance may deter people from engaging in certain legal (or even desirable) online activities because they fear punishment or criminal sanction, and do not trust the legal system to protect their innocence (Penney 2016). Such surveillance "effects" were documented in a survey carried out by the PEN American Center in the immediate aftermath of the Snowden leaks in which they found that writers are engaging in self-censorship as a result (PEN 2013). Further studies have shown a reluctance amongst citizens to engage with politically sensitive topics online, such as a decline in "privacy-sensitive" search terms on Google (Marthews and Tucker 2015), a decline in page views of Wikipedia articles relating to terrorism (Penney 2016), and a "spiral of silence" in surveillance debates on social media (Hampton et al. 2014). As Greenwald claims:

> Merely organizing movements of dissent becomes difficult when the government is watching everything people are doing. But mass surveillance kills dissent in a deeper and more important place as well: in the mind, where the individual trains him- or herself to think only in line with what is expected and demanded.
>
> *(2014, 177–178)*

Furthermore, the Snowden leaks revealed the expansive notion of "target" that has come to be operationalized in such a mode of governance, going far beyond what may be obvious misconduct or wrong-doing. Greenwald points out:

> The perception that invasive surveillance is confined only to a marginalized and deserving group of those "doing wrong" – the bad people – ensures that the majority acquiesces to the abuse of power or even cheers it on. But that view radically misunderstands what goals drive all institutions of authority. "Doing something wrong" in the eyes of such institutions encompasses far more than illegal acts, violent behavior and terrorist plots. It typically extends to meaningful dissent and any genuine challenge. It is the nature of authority to equate dissent with wrongdoing, or at least with a threat.
>
> *(Greenwald 2014, 183)*

The Snowden leaks provided substantial evidence for the ways in which a wide range of politically active citizens are under scrutiny in this ever-expanding threat environment. For example, documents showed that government agencies in both the US and the UK have actively been engaging in the monitoring of political groups with a "watchlist" including international organisations such as Medecins Du Monde (Doctors of the World), UNICEF, Amnesty International and Human Rights Watch, as well as prominent individuals such as Ahmad Muaffaq Zaidan (Al-Jazeera's Pakistan Bureau Chief), Agha Saeed (a former political science professor who advocates for Muslim civil liberties and Palestinians rights), and groups such as Anonymous (Harding 2014; Privacy International and Amnesty International, 2015). State surveillance practices have also extended to the monitoring of politically-interested citizens with programmes such as the one carried out by GCHQ in the aftermath of the "Cablegate" publications which sought to track any visitor to the Wikileaks site by tapping into fibre-optic cables and collecting IP addresses of visitors to the site as well as the search terms used to reach the site (Greenwald and Gallagher 2014).

These disclosures build on previous and continued practices of surveillance of activist groups and dissenting voices. In the UK, revelations of undercover police officers infiltrating a range of activist groups over a longer period of time, including environmental and animal rights activists, have illustrated the invasive tactics used to monitor and suppress protest and dissent (Lubbers 2015). This is alongside other documented forms of managing and containing resistance, tracking activities and intercepting planned actions, whether by corporate agencies or state bodies (cf. Lubbers 2012; Smith and Chamberlain 2015; Uldam 2016). The navigation and circumvention of surveillance is therefore a fully integrated and long-standing tradition in some activist circles (della Porta 1996; Earl 2003; Leistert 2013). However, with the emergence of big data-driven surveillance programmes, regimes of governance and control have increasingly been based on digital infrastructures that facilitate "dataveillance" – a form of continuous surveillance through the use of (meta)data (Raley 2013). These regimes are rooted in the economic logic of "surveillance capitalism" in which accumulation is pursued through the ability to extract, monitor, personalize, and experiment based on the pervasive and continuous recording of digital transactions (Varian 2014; Zuboff 2015). Not only does the entrenchment of this logic within everyday communication technologies cement a fundamentally asymmetrical power relation between activists and those wishing to carry out surveillance on them (Leistert 2012), but the nature of these, often invisible, infrastructures also carries with it central pertinence and significance for activists seeking to challenge existing power relations and mobilize social change. As Lovink and Rossiter (2015) have argued, a politics of the "postdigital" in which the digital has become so omnipresent that it has been pushed to the background and become naturalized, demands of activism to focus on the network architectures at the centre of power in order to pursue genuine social justice and emancipatory ideals.

Anti-surveillance and techno-legal resistance

Efforts to resist data-extractive technologies have taken several forms, particularly in relation to surveillance. As Mann and Ferenbok (2013) have argued, multiple types of "veillance" intersect, undermine and challenge each other in the monitoring of modern societies. Surveillance – veillance in which the viewer is in a position of power over the subject – is often met with efforts to revert or 'equalize' such power. Mann has placed emphasis on the advent of "sousveillance" in this regard, where the subject is gazing back at power "from below", exemplified by technologies such as wearable cameras and other efforts to capture, process, store, recall and transmit human-centered sensory information (Mann 2005, 636). However, as Bakir (2015) points out, modes of resistance to surveillance also include counterveillance and univeillance that speak more to the sabotaging and blocking of surveillance as well as ways of making intelligence services more accountable.

Much resistance to surveillance following the Snowden leaks has centred on these latter strategies – particularly on developing and "mainstreaming" alternative technologies alongside campaigns for tighter policies on the protection of personal data. To start with, forums to provide secure digital infrastructures to activists have proliferated, with "numerous digital rights and internet freedom initiatives seizing the moment to propose new communication methods for activists (and everyday citizens) that are strengthened through encryption". (Aouragh et al. 2015, 213). These have included renewed focus on privacy-enhancing tools such as the TOR browser, the GPG email encryption system and the encrypted phone and text messaging software Signal. An increasing number of websites now support the more secure https protocol rather than the standard http, and a growing number of internet users have downloaded tools such as "https everywhere" that connect to those more secure websites. Privacy guides such as the Electronic Frontier Foundation's "Surveillance Self-Defense" (https://ssd.eff.org/en) and the Tactical Tech Collective's "Security in a Box" (https://tacticaltech.org/projects/security-box) explain the use of privacy-enhancing tools and offer advice on secure online communication. "Crypto-parties" have brought necessary training in such tools to towns and cities worldwide (O'Neill 2015).

Technical solutions to mass data collection have included, furthermore, the development of self-organized communications infrastructures as alternatives to corporate services such as Google and Facebook. Groups such as Riseup.net, Autistici and Sindomino have offered mailing lists, blog platforms and collaborative online workspaces that protect user privacy and are hosted on the groups' own secure servers. Indymedia, arguably the first social media platform, was run by activists in the same manner, and attempts to create other non-commercial and privacy-enhancing social networks have continued. The development of technological alternatives that reinforce autonomous and civil society-based media infrastructure has been a key part of anti-surveillance activism (Hintz and Milan, 2013). Their adoption by activist communities may have grown since the Snowden leaks began but remains limited, so far, as the vast resources available to

large corporate providers and the ease of use of their products – from Gmail to YouTube to Facebook – have meant a far more widespread uptake (Askanius and Uldam 2011; Terranova and Donovan 2013).

However, following the Snowden leaks internet companies have had to address customer concerns regarding data security, too. While they mostly enjoyed friendly relations with, in particular, the US government in pre-Snowden times, divisions between the industry sector and the state emerged after Snowden as criticism of these companies' data practices grew (Wizner, 2017). The confrontation between the FBI and Apple in early 2016 crystallized this new and troubled relation (even if momentarily), in which Apple managed to appear as protector of user interests against state intrusions. The introduction of end-to-end encryption by services such as WhatsApp demonstrated a new trend which aligned, to a degree, with the efforts of non-commercial tech activists. Campaign projects such as "Ranking Digital Rights" (https://rankingdigitalrights.org/) have advanced the focus on corporate policies by, for example, creating an "Accountability Index" that measures company commitment to user privacy and freedom of expression.

While the focus on infrastructure providers and technological development has been prominent, many digital rights campaigns have addressed the state and sought policy reform. In the UK, organisations such as Privacy International, the Open Rights Group, Big Brother Watch, Article 19 and Liberty have regularly issued statements regarding their concerns about mass data collection, have organized public debates and have lobbied legislators. As an immediate response to the Snowden leaks, these groups and others formed a coalition – Don't Spy On Us – which combined some of this advocacy work towards a common campaign. Their voice was significant in the specialized discourses around, for example, the Investigatory Powers Act – the main post-Snowden piece of UK legislative reform. They have formulated fundamental critiques of surveillance practices, but they have also, increasingly, been recognized as a legitimate participant in policy debates that holds relevant expertise. As one anti-surveillance campaigner noted: "Previously NGOs would have fought just to kill a new law and probably been unsuccessful in doing so; now they can say: here's how we can genuinely improve it and have a proper conversation with the Home Office" (quoted in Hintz and Brown 2017).

Litigation has emerged as a key strategy of policy advocacy. Campaign organisations such as Privacy International, Liberty and Amnesty International challenged GCHQ's data collection practices at the Investigatory Powers Tribunal (IPT) which decided that some of the agency's activities were unlawful. Others, such as the Open Rights Group, Big Brother Watch and Human Rights Watch brought cases against the British government before the European Court of Human Rights and the European Court of Justice. While the results of legal challenges have been mixed, they have forced governments to admit to previously secret practices and have thereby opened up avenues for policy reform (Hintz and Brown 2017).

At the intersection between policy and technology, civil society activists have also contributed to the work of institutions that define and regulate the standards and protocols of digital communication. In some of these bodies, such as the

Internet Engineering Task Force (IETF), they participate in individual capacity and based on their personal expertise, next to experts from industry and government. In others, such as the Internet Corporation for Assigned Names and Numbers (ICANN), they form specific caucuses, for example the Non-commercial User Constituency (NCUC). As technical standards and protocols typically allow some actions and disallow others, and enable some uses and restrict others, their development constitutes a latent and invisible form of policymaking and therefore places standards organisations in both a highly influential and slightly obscure position (cf., DeNardis 2009; Lessig 1999). In response to the Snowden leaks, several of these bodies started to address the vulnerabilities exposed in the revelations by setting up working groups, developing proposals on how to incorporate privacy in standards, and, in some cases, agreeing that these concerns should become a priority of standards development (Rogers and Eden 2017).

Digital rights activists and civil society-based technological developers have been influential in all these venues. Yet their efforts have largely remained within a specialized discourse and a constituency of experts. Our goal with this research was to explore to what extent activists concerned with other social justice issues have engaged with these agendas, and whether there is scope for linking these (possibly) divergent concerns.

Resistance to datafication amongst political activists

In the rest of this chapter, we therefore explore the extent to which such resistance to digital surveillance features in broader activist practices and how concerns with data are understood. This research is based on a number of semi-structured interviews carried out with political activists in the UK as part of the larger project "Digital Citizenship and Surveillance Society: UK State-Media-Citizen Relations After the Snowden Leaks". These interviews were conducted with a range of political activists, both from big NGOs as well as smaller community and grassroots organisations based in the UK, that were not specifically engaged with digital rights or technology activism, and individuals within those groups who were not specifically responsible for technical infrastructures of communication. These groups were chosen on the basis of having a more or less adversarial relationship with the state, covering a range of causes, and predominantly out of an existing network of contacts. They therefore cover a relatively wide spectrum of civil society activity. The sample consisted of 11 interviews (see Table 9.1) carried out in person (8) or on Skype (3) during March–June 2015, lasting on average 60 minutes and focused on the following themes: a) understanding and experience of surveillance; b) knowledge and opinions of the Snowden leaks; c) attitudes towards state surveillance; d) online behaviour and practices; e) changes and responses to the Snowden leaks.

In the context of the above discussion, this chapter is particularly concerned with the extent to which resistance to digital surveillance features in activist practices and agendas and how data politics more broadly is understood. We extracted prominent themes from our interviews around these issues, based on a thematic analysis that

TABLE 9.1 List of Interviews

Organization	*Orientation*
Global Justice Now (GJN)	Economic justice
Campaign Against Arms Trade (CAAT)	Anti-arms
CAGE	Anti-discrimination
Muslim Association of Britain (MAB)	Community integration
Greenpeace	Environmentalism
Stop the War Coalition (STWC)	Anti-war
Muslim Council of Wales (MCoW)	Community integration
Trade Union Congress (TUC)	Workers' rights
Anti-fracking activist	Environmentalism
ACORN	Community organizing (housing)
People's Assembly Against Austerity (PAAA)	Anti-austerity

focused on understandings of surveillance, uses of encryption software, changes in communication practices following the Snowden leaks, and attitudes towards digital rights advocacy. Below we outline key themes emerging from our interviews in relation to how anti-surveillance and data politics is situated in activist practices. In the first part we discuss general understandings of surveillance, data collection, and responses to the Snowden leaks. In the second part we move on to discuss how resistance to mass data collection in terms of encryption and advocacy around digital rights is understood and practiced amongst the activists we interviewed.

Responses to Snowden

To start with, the interviews demonstrated that the issue of state surveillance is very familiar amongst political activists in the UK, particularly due to a troublesome history of police infiltration into activist groups. Many of the activists we spoke with had either direct experiences of police infiltrating groups they were part of or they knew someone who had experienced infiltration. Digital surveillance and big data surveillance of the kind revealed in the Snowden leaks was less prominent and salient in initial descriptions of surveillance. However, many of the activists we interviewed expressed a general awareness and expectation that these activities are going on, from either corporations or state, or a combination of both. Several activists pointed to specific experiences that might demonstrate the monitoring of online activities:

> I think there's been instances where the police have turned up to our meetings or rung ahead of venues we've been using and warned the venues not to allow us to have a meeting (. . .) they're obviously keeping tabs on our Facebook activities but then that's public so you totally expect that. Similarly with Twitter . . . we're pretty sure that there's a police presence on something called Basecamp which is where we organize online.
>
> *(Anti-fracking activist)*

What was revealed in the Snowden leaks, therefore, came as little surprise to the majority of our interviewees although the scale of the surveillance programmes revealed in the documents did exceed expectation for many activists:

> I think, kind of like most people my impression is there has been a hell of a lot more going on than anyone has known about. The capabilities of the security services are much greater than anyone suspected but there is much less political and judicial oversight of this, and indeed some of this is done on a dubious legal basis.
>
> *(TUC activist)*

The lack of surprise, or the widespread expectation, of what the Snowden leaks revealed therefore also muted any direct reaction to the Snowden leaks amongst most of the activists we interviewed. With the exception of Greenpeace who reviewed and revised their communication infrastructure as an immediate and direct result of the Snowden leaks, our interviewees expressed little, if any, direct response to the revelations.[4] Rather, awareness and continued negotiation with the realities of surveillance has developed over time and the Snowden leaks fit into this longer-term consciousness instead of being transformative in and of themselves:

> I think it's about being always aware of the general threat. I don't think in fact that Snowden in particular has had an impact on a single aspect of how we work . . . In a sense he confirmed what was the sort of thing people suspected was happening anyway, but I don't think that revelation has changed anything we do.
>
> *(CAAT activist)*

Of course, this does not mean that precautions are not taken against digital surveillance as part of activist practice. Some of the people we interviewed spoke of tactics employed to circumvent different forms of digital surveillance, such as using anonymisation tools (e.g. a VPN) for researching targets, preferring face-to-face meetings for organizing actions, and using encrypted emails for sharing personal data. This also highlights how circumvention of data collection is more prominent for particular kinds of activities (e.g. internal organizational use). Overwhelmingly, however, our interviews illustrate the extent to which the dependence on digital communications, and mainstream social media in particular, for pursuing activist agendas undermines efforts to actively circumvent or resist data-driven surveillance. Activist groups use digital infrastructures that are subject to large-scale data collection for several aspects of their activities, including general awareness-raising, advocacy, mobilizing, organizing and expanding their actions and membership base, using programmes and tools integrated into social media interfaces. They do so because of the perceived reach that social media platforms afford and because activists themselves rely on the "datafication"

of social relations in order to collect data and extend networks of connections, both for organization and mobilization of activities:

> [NationBuilder] is a programme which is designed for campaign organisations. Obama used it in his campaign. Labour are using it. It basically integrates your website with a database and social media as well so sucks in social media profiles out of Facebook and Twitter and things like that.
>
> *(ACORN activist)*

> You start off by setting up a Facebook event and then the activists learn tools like the invite all app where you don't have to keep on inviting individual friends, it invites 500 at a time. So we will then spread that all around people so then you can drive the invites up to 5 or 6 thousand very quickly.
>
> *(PAAA activist)*

Such dependency on this kind of digital infrastructure in conjunction with a general awareness that communication is being monitored and stored in turn manifests itself by forms of self-regulating online behaviour. Despite their widespread use of mainstream platforms, activists noted they are cautious about not saying anything "too controversial" on social media, or choose to withdraw entirely from using social media to discuss politics:

> My advice to our people, our community, is just be careful before saying anything, before making a statement . . . and think about it, what the repercussions would be and how it could be misconstrued. So prevention is better.
>
> *(MCoW activist)*

> It can get picked up and used in a court or, partly in a court case or possibly liable. I think people are worried about liable.
>
> *(STWC activist)*

These types of concerns speak partly to the "chilling effect" mentioned above in which some online activities and communication are deterred out of a fear of the repercussions and mistrust towards the system.

Resisting data collection

Despite such concerns being expressed, the active circumvention of data-extracting technologies such as widespread uptake of encryption or anonymisation tools remained limited to just a few of the groups we interviewed, with Greenpeace expressing the most extensive and comprehensive secure communication infrastructure. Predominantly, the activists we interviewed did not use encryption or anonymisation tools as an integrated part of their communication practices. In reasoning this, we can see a number of themes emerge. Firstly, several interviewees

spoke of a perceived "lack of knowledge", insufficient technical ability and not being able to "afford" to implement alternative communication practices. These kinds of perceptions are often also combined with notions of convenience in which mainstream platforms are favoured for their familiarity and ease of use:

> We just want ease of access to be honest. Actually, I can send an email to a few thousand people and do a few other things and I don't need to spend days or weeks actually learning how to do it because I'm not very technically minded.
> *(ACORN activist)*

The question of convenience is linked to a second significant theme that emerged on this topic. Activists feel that using encryption strides against their ambitions of being an "open" and "inclusive" group or organization. Several of our interviewees emphasized the transparent nature of their activities, including also the legality of their tactics, and their wish to be a "public" movement. In positioning their response in this way, we can identify an important perception of encryption as being linked to "hidden" practices or "exclusive" forms of communication. In contrast to understanding encryption according to its established purpose as a means of security and protection, and as an enabler of both privacy and freedom of expression (Kaye 2015), the strong role of a popular "nothing to hide" discourse is evident even among activists. A number of interviewees understood such tools as contradicting or undermining their self-identification:

> We've got nothing to hide, we're not doing anything illegal and we're not doing anything that's not defendable. So you know . . . if the security services want to challenge what we're doing then we'll have that debate out in public. And anyway, I suppose at the back of our minds is that it probably wouldn't work anyway is my guess. Without spending huge amounts of time or resources.
> *(STWC activist)*

> We're having to campaign all of the time, we're not secret organisations, or organisations of tight-knit groups of people campaigning together. We are mass movements, and we are open. For us social media is great because it makes communication easy and of course we know people look at social media but our messages are not hard to get.
> *(TUC activist)*

The point here is not the choice of tactics that these groups use. Rather, the attitude expressed here demonstrates that privacy-enhancing technology is seen to be pertinent to only a particular strand of political activism and directly undermines another. Indeed, there was a prevalent sentiment in several of our interviews that being part of "mainstream" groups reduced the need for concern with digital surveillance practices. That is, resisting or circumventing data collection as an activist practice is predominantly confined to those engaging in "radical" political activism.

Consequently, this might also deter those "in the middle" from becoming more "radical", making "people more cautious" (ACORN activist), and thereby keep the mainstream "in check". This sentiment is reasoned not just in terms of the legality of tactics that different activist groups employ, but also in terms of the perception of their own influence and the extent to which they see themselves as adversarial to the state (our sample includes a variation of activists in this regard). In this sense, only activists who understand themselves as being sufficiently of interest to the state feel the need to concern themselves with data-driven surveillance as an issue or integrate secure technologies into their practices.

Such perceptions also extend to activists' engagement with advocacy on legislation relating to privacy or digital rights issues more broadly. Although solidarity and support of the cause was expressed across the board, most activists we interviewed did not see themselves or the organisations and groups they are part of as being actively engaged with issues relating to digital rights, such as privacy or data protection. Rather, despite mentions of some informal links with organizations such as Privacy International and Statewatch[5], most of the activists we interviewed made a distinction between their own activist work and that of technology activists and digital rights groups:

> Some people focus on things like surveillance and some people focus on the workplace, some people do community things.
>
> *(ACORN activist)*

> I think there are organisations that are doing that work already and it's for us to be knowledgeable and a bit of a step ahead of the game, but I don't think it's for us to campaign on surveillance.
>
> *(Anti-fracking activist)*

Despite a general critique of mass data collection, resisting it actively does not feature in activists' own agendas and is instead "out-sourced" to expert communities. In this sense also, resistance to data collection was not seen as providing a base for a broader movement, but rather an issue in which you need to "specialize" (PAAA activist).

"Data justice" and the bridging of activism(s)

Our interviews with activists illustrate that a general awareness and expectation of surveillance is prevalent amongst activist communities in the UK, but concerns with data-driven surveillance of the kind revealed in the Snowden leaks remain somewhat marginalized in activist perceptions and practices. Rather, the entrenched dependency on mainstream communication platforms that are predominantly insecure provide an environment for activist practices in which it is seen as difficult and problematic to engage in resistance to data collection either through technological means or in terms of protest and advocacy for

greater privacy and data protection. More generally, we can identify a "disconnect" between concerns with data-driven surveillance and other (broader) social justice concerns.

How, then, might we address this disconnect? Aouragh et al. (2015) argue that the "division of labour" between what they label "tech justice" and "social justice" activists emerges partly from the socio-technical practices that have been advanced in secure communication campaigns in which there is a distinct user-developer dichotomy that places the onus on the (individual) "user" to protect themselves (identifying risks using "threat modeling") with tools provided by the "developer". Similarly, Kazansky (2016) found, based on her experience with providing information security training for human rights activists, that training is often designed towards the individual user rather than as a collective project that considers the enabling social structures needed for secure communication to become an integrated activist practice. This speaks to the shortcomings identified by Ruppert, Isin and Bigo (2017), with the common atomism prevalent in views on data politics and the onus on immediacy that pervades responses. That is, the view of the internet as addressed to atomized individuals who then need to protect themselves against the immediacy of a threat engendered. This view, they argue, is based on the ontological premise of "hyper-individualism" in which the addressee is "the atomized subject whose data is individualized rather than understood as a product of collective relations with other subjects and technologies" (Ruppert, Isin and Bigo 2017, 3) Policy reform advocacy, meanwhile, does not address individual users but, nevertheless, the specific audience of policymakers and thereby erects different boundaries, based on issue-specific expertise and discourse (Hintz and Brown 2017).

Such approaches, Aouragh et al. contend, configure modes of delegation that actually come to negate possibilities for overlaps between different justice claims and reproduce "a perhaps unintended hierarchy based on traditional models of production" (2015, 216). Drawing on their research with "tech justice" activists, they therefore argue for connecting security engineers with the language of collective action within a political project and, more broadly, for dissolving the perceived divisions of justice claims that persist between these activist camps.

Building on this ambition, we want to further advance the debate based on our research with "social justice" activists by suggesting a broader framework that may allow us to develop a more integrated understanding of data collection in relation to social justice agendas. As outlined above, the terms upon which resistance to surveillance has predominantly been approached have placed data debates within the parameters of particular expert communities, namely technology activists and digital rights groups. This techno-legal framing of resistance, although partly dictated by the activist opportunity structures currently available, limits our understanding of the implications of these data-driven practices that underpin contemporary surveillance and dilutes their politicized nature. The consequences of this limitation include, for example, a relatively uncritical perspective among digital rights advocacy communities on "targeted" surveillance which is often seen as a benign alternative to

indiscriminate "mass" surveillance but abstracts from the experiences of minority communities and political activists as typically targeted groups (Gürses, Kundnani and Van Hoboken 2016). Further, this limited perspective may lead to a perception of industry surveillance as largely politically benevolent and the turn to the tech companies of Silicon Valley as our "protectors" in the counter-surveillance struggle, armed with PR-friendly encryption tools. Moreover, a techno-legal framing of the issue risks masking the struggles through which people come to be governed by data (Ruppert, Isin and Bigo 2017). As Gürses, Kundnani and Van Hoboken (2016) suggest, these problematic positions point to the need for a political analysis as our starting point for countering the systems of data collection and use that have been developed; one that simultaneously broadens the discussion beyond the narrow confines of techno-legal parameters and speaks to the concerns of activists across technology and social justice camps.

As part of such an analysis, we advance the notion of "data justice" as a way to highlight the place of data-driven surveillance, and related big data decision-making and governance, in conceptions of social justice. Whilst recognizing the procedural inference in the term "justice", by data justice we are referring to the implications that data-driven processes at the core of surveillance capitalism have for the pursuit of substantive social and economic justice claims. This, we suggest, encompasses both the targeting of surveillance against activists leading to repression, self-censorship and chilling-effects in the organization, mobilization, and pursuit of social justice as well as the role of data collection in (new) forms of governance that shape society in line with particular political and economic agendas. As Andrejevic (2015) has outlined, the nature of the surveillance programmes revealed in the Snowden leaks are intimately linked to a system of economics and a state-corporate interest not necessarily in individual people, but in detecting and predicting patterns, profiling and sorting groups. Big data surveillance brings up issues not just of privacy, but also of social sorting and preemption (Lyon 2014) and is generative of new power relations and politics (Ruppert, Isin and Bigo 2017). Although much more difficult to ascertain in concrete terms, this has significant implications for people's lives and the society they will live in. Data justice as a framework is intended to guide a research trajectory and types of activity that bring out and underscore this politics of data and the implications of these practices for substantive social justice claims. This is obviously a bigger task beyond our current scope, but here we can highlight some questions that have already planted the seeds for further illumination and advancement of our understanding of resistance to datafication in this regard.

Whilst "data justice" as a concept and framework is still in nascent form, different interpretations are being advanced that share a concern with outlining data in relation to structural inequality and social (in)justice (Heeks 2017; Heeks and Renken 2016; Hintz, Dencik and Wahl-Jorgensen 2018; Newman 2015; Taylor 2017). Grassroots groups and social justice campaigns have started to apply this more comprehensive approach to datafication and, in some cases, have done so within a "data justice" framework. For example, the Detroit Digital

Justice Coalition (http://detroitdjc.org) has worked with local residents in identifying potential social harms that may emerge through the collection of citizen data by public institutions. In particular, they are concerned with the criminalization and surveillance of low-income communities, people of colour and other targeted groups. As a result, they have developed a set of guidelines for equitable practices in collecting, disseminating and using data in relation to social and historical context. The US/Canadian Environmental Data & Governance Initiative, EDGI (https://envirodatagov.org/), has preserved vulnerable scientific data in the aftermath of the US election of Trump in 2016 and, in the process, has developed a deeper understanding of the politics, generation, ownership and uses of environmental data. Their perspective on "environmental data justice" "brings together the concerns of the emergent area of data justice with the long-standing principles of environmental justice" (https://envirodatagov.org/towards-edj-statement/). The local administration in Barcelona, meanwhile, has been actively developing alternative infrastructures, with an emphasis on decentralized technologies that are designed for more citizen-led and participatory platforms and where ownership of data belongs to the citizens. Such ideas are also prevalent in the growing "platform cooperativism" movement that sets out to challenge the dominance of contemporary platform capitalism in order to create a fairer future of work in a digital economy by building on the values of cooperativism.

These sorts of initiatives speak partly to the framework we are proposing here by reframing data debates to consider how digital infrastructures and data-driven processes have implications for broader society beyond individual privacy. We want to further progress this agenda by suggesting that "data justice" can provide a conceptual foundation for exploring how mass data collection implicates different understandings of social justice as well as a potential action-building tool for addressing such implications. This requires us to further examine the ideological basis of data-driven processes, situating this form of governance within a political agenda that extends to particular conceptions of society and the demarcation of "good" and "bad" citizens. Furthermore, it leads us to scrutinize the interests and power relations at play in 'datafied' societies that enfranchise some and disenfranchise others, highlighting also forms of exclusion and discrimination. Moreover, it requires us to stipulate how society is and ought to be organized in relation to digital infrastructures – on social, political, economic, cultural and ecological terms – that can consider and develop the meaning of justice in this context. This includes questions of how to think about notions such as security, autonomy, dignity, fairness and sustainability in a data-driven society and make us ask what, for example, the implications are for community cohesion and discrimination; for welfare and inequality; for workers' rights; or for the environment, for poverty, and for conflict. Most importantly, advancing this agenda transforms data-driven processes from a special-interest "issue" into a core dimension of social, political, cultural, ecological and economic justice, and thus responds to the central position of data in contemporary capitalism.

By advancing the framework of "data justice" our point is to illustrate how the relationship between political activism and surveillance is not one in which activists are only at risk for expressing dissent, but one in which the very infrastructures of surveillance (big data) have direct consequences for the social justice claims they are seeking to make. Data justice integrates the concerns with the collection, use and analysis of data with activists' agendas, not just to protect themselves, but also to achieve the social change they want to make. As such, this may offer an opportunity to bridge the current "disconnect" we have found in anti-surveillance resistance and provide resources for a political and social movement that can engage with data debates beyond techno-legal solutionism. It may also contribute to the call for a conception of data subjects more in relation to rights-claiming citizens (Ruppert, Isin and Bigo 2017). This, we would argue, is urgently needed in the shift towards data-driven forms of governance rooted in surveillance capitalism.

Conclusion

The need to engage with data politics in a way that situates concerns with data-driven processes within a broader framework of social justice is becoming increasingly prevalent, and is being approached from a number of different angles.[6] In this chapter we have addressed this topic by looking specifically at attitudes and practices pertaining to resisting data-driven surveillance. The Snowden leaks constitute an important moment for exploring these questions as they provided substantial evidence for the extent of the collection and use of digitally-generated data (or big data) and illustrated the intricate relationship between the infrastructures of our everyday technologies and emerging forms of governance and control. Pertinent debate, activity and advocacy has flourished in response to the Snowden leaks, opening up opportunities for many existing technology- and digital rights-concerned communities to mobilize, expand and influence political processes and social attitudes. However, due to the dominant political culture and opportunity structures available to active participants in the resistance against surveillance, debates on data collection and use, and critical engagement with data politics more broadly, have struggled to move beyond the participation of particular expert communities. A concern with digital surveillance, in this context, has come to be viewed as a "specialist" issue in which achieving "tech justice" is predominantly centred on technical and legal solutions relating to privacy and data protection. We have seen this in our research on attitudes and practices amongst political activists engaged in broader social justice issues, from environmentalism to labour justice to anti-discrimination, who have predominantly come to view data-driven surveillance as an issue that does not substantially feature on their agenda.

Rather, what emerges in the broader ecology of civil society pursuits of justice, is a kind of "disconnect" between those concerned with technology issues and those concerned with social justice issues as two separate camps. Of course, we recognize that this comes partly from the necessity to set priorities and focus on particular

topics when activist energies and resources are frequently limited. However, we argue that the nature of surveillance revealed in the Snowden leaks speaks to an urgent need to broaden the parameters for how data collection and use has been understood and discussed that implicates activists across the tech and social justice camps, collectively. The ability to monitor, record and store digital transactions on a massive scale creates an environment that substantially limits the possibilities for dissent and protest, whether through self-censorship, chilling-effects or active repression. Moreover, however, it constitutes a form of governance that is rooted in and simultaneously advances particular social, economic, and political agendas that enfranchise some whilst disenfranchising others, and prioritizes certain ways of organizing society at the expense of others.

By introducing the notion of "data justice" in this chapter we want to contribute to the shift and broadening of our understanding of the role of data-driven processes in contemporary society. By advancing data justice as a framework for debate and research, we want to set the parameters for a discussion on datafication that can illuminate the implications for social justice, both in terms of the conditions for communicating autonomously and practicing dissent as well as the social and economic (in)justices that are produced by this form of governance (and, therefore, what might be the possible alternatives). Referring to "data justice" recognizes the political economy of the system that underpins the possibilities for extensive data collection and use, whilst drawing attention to the political agenda that is driving its implementation. This, we argue, comes to impact on political activists and their pursuits of social justice in significant ways and provides an impetus for a broad collective movement to engage in pertinent data-related debates. Such a collective approach is needed, we suggest, in light of a shift to surveillance capitalism in which the collection, use and analysis of our data increasingly comes to shape the opportunities and possibilities available to us and the kind of society we live in.

Notes

The research for this chapter was funded by the Economic and Social Research Council as part of the project "Digital Citizenship and Surveillance Society: UK State-Media-Citizen relations after the Snowden leaks".

1 Note that a version of this was originally published in *Big Data & Society*.
2 The chapter is based on research carried out for the collaborative research project "Digital Citizenship and Surveillance Society: UK State-Media-Citizen relations after the Snowden leaks" at Cardiff University funded by the Economic and Social Research Council, the first comprehensive review of the implications of the Snowden revelations from a UK perspective.
3 Details of the revelations can be found at The Snowden Archive: http://www.cjfe.org/snowden
4 We were later informed that CAGE has also significantly changed their communication infrastructure, but this development happened after our interview period.
5 Interestingly, also, CAGE participated for the first time in the large hacker convention Chaos Communication Congress in December 2015.
6 See, for example, the 'Data Justice Conference' that took place in Cardiff in the United Kingdom in May 2018: https://datajusticelab.org/data-justice-conference/

References

Andrejevic, M. 2015. Keynote plenary at the conference 'Surveillance and Citizenship', Cardiff, 19 June.

Aouragh, M., Gürses, S., Rocha, J. and Snelting, F. 2015. Let's first get things done! On division of labour and techno-political practices of delegation in times of crisis. *The Fibreculture Journal*, 26: 208–235.

Askanius, T. and Uldam, J. 2011. Online social media for radical politics: climate change activism on YouTube. *International Journal of Electronic Governance (IJEG)*, 4(1/2): 69–84.

Bakir, V. 2015. Veillant panoptic assemblage: Mutual watching and resistance to mass surveillance after Snowden. *Media and Communication*, 3(3): 12–25.

della Porta, D. 1996. Social Movements and the State: Thoughts on the Policing of Protest. In D. McAdam, J. McCarthy, and M. N. Zald (eds.), *Comparative Perspectives on Social Movements. Political Opportunities, Mobilizing Structures, and Cultural Framing*. Cambridge/New York: Cambridge University Press, 62–92.

DeNardis, L. 2009. *Protocol Politics: The Globalization of Internet Governance*. Cambridge: MIT Press.

Earl, J. 2003 Tanks, tear gas, and taxes: Toward a theory of movement repression, *Sociological Theory*, 21(1): 44–68.

Greenwald, G. 2014. *No Place to Hide: Edward Snowden, the NSA and the surveillance state*. London: Hamish Hamilton.

Greenwald, G. & Gallagher, R. 2014. Snowden documents reveal covert surveillance and pressure tactics aimed at WikiLeaks and its supporters, The Intercept, February 18, available at: https://theintercept.com/2014/02/18/snowden-docs-reveal-covert-surveillance-and-pressure-tactics-aimed-at-wikileaks-and-its-supporters/ (last accessed March 7th, 2016).

Gürses, S., Kundnani, A. and Van Hoboken, J. 2016. Crypto and empire: the contradictions of counter-surveillance advocacy. *Media, Culture & Society*, DOI: 10.1177/0163443716643006.

Hampton, K.N., Rainie, L., Lu, W., Dwyer, M., Shin, I., and Purcell, K. 2014. 'Social media and the 'spiral of silence''. Pew Research Center, Washington, DC. Retrieved from http://www.pewinternet.org/files/2014/08/PI_Social-networks-and-debate_082614.pdf.

Harding, L. 2014. Edward Snowden: US government spied on human rights workers, *The Guardian*, April 8.

Heeks, R. and Renken, J. 2016. Data justice for development: what would it mean? *Information Development*. Available at: https://doi.org/10.1177/0266666916678282.

Heeks, R. 2017. A structural model and manifesto for data justice for international development. *Development Informatics Working Paper Series*, No. 69.

Hintz, A. and Brown, I. 2017. Enabling digital citizenship? The reshaping of surveillance policy after Snowden. *International Journal of Communication*, 11: 782-801.

Hintz, A., Dencik, L. and Wahl-Jorgensen, K. 2018. *Digital Citizenship in a Datafied Society*. Cambridge: Polity Press.

Hintz, A. and Milan, S. 2013. Networked collective action and the institutionalised policy debate: Bringing cyberactivism to the policy arena? *Policy & Internet*, 5(1): 7–26.

Kaye, D. 2015. Report on encryption, anonymity, and the human rights framework. Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. 22 May 2015. Available at: http://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session29/Documents/A.HRC.29.32_AEV.doc.

Kazansky, B. 2016. Digital security in context: Learning how human rights defenders adopt digital security practices. Project report. Available at: https://secresearch.tacticaltech.org/media/pages/pdfs/original/DigitalSecurityInContext.pdf?1459444650 (last accessed 21 May 2016).

Kitchin, R. 2014. *The Data Revolution*. London: Sage.
Leistert, O. 2012. Resistance against cyber-surveillance within social movements and how surveillance adapts, *Surveillance & Society*, 9(4): 441–456.
Leistert, O. 2013. *From Protest to Surveillance – The Political Rationality of Mobile Media*. Frankfurt Am Main: Peter Lang.
Lessig, L. 1999. *Code and other Laws of Cyberspace*. New York: Basic Books.
Lovink, G and Rossiter, N. 2015 Network Cultures and the Architecture of Decision. In Dencik, L. & Leistert, O. (eds.) *Critical Perspectives on Social Media and Protest: Between Control and Emancipation*. London: Rowman & Littlefield International: 219–232.
Lubbers, E. 2012. *Secret Manoeuvers in the Dark: Corporate and Police Spying on Activists*. London: Pluto Press.
Lubbers, E. 2015. Undercover Research: Corporate and police spying on activists. An introduction to activist intelligence as a new field of surveillance. *Surveillance & Society*, 13(3/4): 338–353.
Lyon, D. 2001. *Surveillance Society: Monitoring Everyday Life*. Buckingham and Philadelphia: Open University Press.
Lyon, D. 2014. Survellance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data & Society*, July–December: 1–13.
Lyon, D. 2015. *Surveillance After Snowden*. Cambridge and Malden, MA: Polity Press.
Mann S 2005. Sousveillance and cyberglogs: A 30-year empirical voyage through ethical, legal and policy issues. *Presence: Teleoperators and Virtual Environments*, 14(6): 625–646.
Mann, S. and Ferenbok, J. (2013) New Media and the power politics of sousveillance in a surveillance-dominated world. *Surveillance & Society*, 11(1/2): 18–34.
Marthews, A. and Tucker, C. 2015. Government surveillance and internet search behaviour. *Available at SSRN:* http://ssrn.com/abstract=2412564.
Massumi, B. 2015. *Ontopower: War, Powers, and the State of Perception*. Durham and London: Duke University Press.
Newman, N. (2015) Data Justice: Taking on Big Data as an Economic Justice Issue. Report. Available at: http://www.datajustice.org/blog/data-justice-report-taking-big-data-economic-justice-issue
O'Neill, P. H. 2015. The state of encryption tools, 2 years after Snowden leaks, *The Daily Dot*, 20 June. Available at: http://www.dailydot.com/layer8/encryption-since-snowden-trending-up/ (last accessed 4 September 2016).
PEN. 2013. *Chilling Effects: NSA Surveillance Drives U.S. Writers to Self-Censor*. New York: PEN American Center, available at: http://www.pen.org/sites/default/files/Chilling%20Effects_PEN%20American.pdf (last accessed 7 March 2016).
Penney, J. 2016. Chilling effects: Online surveillance Wikipedia use. *Berkeley Technology Law Journal*, Available at SSRN: http://ssrn.com/abstract=2769645.
Privacy International & Amnesty International. 2015. Two years after Snowden: Protecting human rights in an age of mass surveillance. Report, available at: https://www.privacyinternational.org/sites/default/files/Two%20Years%20After%20Snowden_Final%20Report_EN_0.pdf (last accessed 7 March 2016).
Raley, R. 2013. Dataveillance and Countervailance. In: Gitelman, L. (ed.) *'Raw Data' is an Oxymoron*. Cambridge, MA: MIT Press: 121–146.
Rogers, M. and Eden, G. 2017. The Snowden disclosures, technical standards and the making of surveillance infrastructures. *International Journal of Communication*, 11: 802–823.
Isin, E. and Ruppert, E. 2015. *Becoming Digital Citizens*. Lanham: Rowman & Littlefield.
Ruppert, E., Isin, E. and Bigo, D. 2017. Data politics. *Big Data & Society*. July– December: 1–7.
Smith, D. and Chamberlain, P. 2015. *Blacklisted: The Secret War Between Big Business and Union Activists*. Oxford: New Internationalist.

Taylor, L. 2017. What is data justice? The case for connecting digital rights and freedoms. *Big Data & Society*, 4(2).

Terranova, T. and Donovan, J. 2013. "Occupy Social Networks: The Paradoxes of Corporate Social Media for Networked Social Movements." In Lovink, G. and Rash, M. (eds.) *Unlike Us Reader: Social Media Monopolies and Their Alternatives*. Amsterdam: Institute of Network Cultures, 296–311.

Uldam, J. 2016. Corporate management of visibility and the fantasy of the post-political: Social media and surveillance, *New Media & Society*, 18(2): 201–219.

Varian, H. R. 2014. Beyond Big Data, *Business Economics*, 49(1): 27–31.

Wizner, B. 2017. What changed after Snowden? A US perspective. *International Joutnal of Communication*, 11: 897–901.

Zuboff, S. 2015. Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30: 75–89

10

THESES ON AUTOMATION AND LABOUR

Brett Neilson and Ned Rossiter

Automation has already happened

Never mind the future. Automation has already ripped through the past. From medieval robots to the fake chess-playing device that appears in the first of Walter Benjamin's (1969) "Theses on the Philosophy of History," automata have troubled "paired ideas about life and death, nature and manufacture, foreign and familiar" (Truitt 2015, 1). As Adelheid Voskuhl notes in *Androids in the Enlightenment* (2013), only in the age of industrial factory production does automation begin to instil the fear that modern selves and societies have become indistinguishable from machines. Andrew Ure, Charles Babbage, and Karl Marx were the prophets of this anxiety. From the spinning jenny to the assembly line of car manufacturing, automation tied industrial modernity to the experience and conditions of labour. Global labour history has taught us that capitalism appeared as much in the plantation as in the factory, and that the worker is as much a slave as a freely contracted individual. We learn a similar lesson by considering the power of the machine. In the "Fragment on Machines," Marx (1973, 693) writes:

> The science which compels the inanimate limbs of the machinery, by their construction, to act purposefully, as an automaton, does not exist in the worker's consciousness, but rather acts upon him through the machine as an alien power, as the power of the machine itself.

Here lies a central paradox that repeats itself in studies of automation and labour. Automation reduces workers to "conscious linkages" of the machine, making labour "a mere abstraction of activity," but the machine in so far as it comprises an *automatic system* becomes "living active machinery," a "mighty organism" (692) endowed "with consciousness and a will" to minimize human resistance (Marx 1977, 527).

If, as Ben Trott (2017) advises, we read Marx's "Fragment on Machines" as science fiction, we need to take seriously the suggestion that automation registers capital's ability to absorb and valorize social knowledge and social life, to the point that accumulation no longer rests exclusively on the sale and purchase of labour-power. Now as before, the fear is that automation makes labour redundant. From the time of the industrial revolution, debate has raged about the implications of automation for employment. Adam Smith (1776) linked machinery to the division of labour and emphasized its labour saving capacity. Jean-Baptiste Say (1803) was confident about the ability of markets to adjust to the introduction of automated machinery. In his 1817 text *Principles of Political Economy and Taxation*, David Ricardo (1951, 392) argued that the labour market would adjust to automation but recognized that concerns about job displacement were "conformable to the correct principles of political economy." Thomas Malthus (1836) maintained that demand dynamics spurred by cheaper production techniques would offset employment loss due to automation. Despite historical change, these positions tend to repeat themselves whenever the emergence of new automated technologies raises the prospect that capital might dispense with labour.

Let us take two twentieth century examples, written in the decades when the introduction of the computer was first beginning to change industrial labour processes. In *Efficiency, Equality and the Ownership of Property*, James E. Meade (1964, 25–26) worried that "new labour required with the new automated machines" would be "less than the growth of the labour force plus the labour made redundant by the scrapping of physically worn-out old machinery." By contrast, Wasily Leontief and Faye Duchin (1986) emphasized the possibility for new employment to arise in the design, operation, and maintenance of automated technologies. In *The Future Impact of Automation on Workers*, they utilized the input-output model of economic analysis to forecast scenarios for the impact of computer-based automation across different sectors of the U.S. economy through to the year 2000. Charting ripple effects across industries by generating coefficients from a series of matrices, they concluded that while automation will have negative employment effects in sectors such as manufacturing, office work, health, and education, there would be inverse effects in sectors such as robotics. In reality, in the years between 1986 and 2000, the effects of computer-based automation intertwined with those of offshoring domestic production from the U.S. and other established industrial nations. Notably this same period corresponded to the emergence of East Asia, and particularly China, as a site of computer hardware manufacture. In considering the effects of automation on labour, we cannot afford to approach, as did Marx when he made assumptions to undergird arithmetic proof of his arguments, "the whole world of trade as one nation" (Marx 1977, 727).

The tendency of automation to play differently across industries, occupations, and nation-states must stay in view today when the integration and extension of digital technologies is again at the centre of a defining transformation of life, economy, and society. Over recent years, there has been a consistent stream of news reports and policy documents forecasting the displacement of labour not only

in manufacturing but also in white-collar industries (see, for instance, Frey and Osborne 2013, Hoskins 2016, Méda 2016, Regalado 2012). Going beyond Harry Braverman's warnings about deskilling and the separation of manual and mental work in *Labor and Monopoly Capital* (1974), the claim is that the economy will not reabsorb these positions according to the familiar model by which new jobs emerge in science-based and technological sectors. In other words, the structural ascent of workers into higher level tertiary industries and economies hit a peak over the last thirty years and is now facing gradual, and in some sectors quite dramatic, termination. This has implications beyond the sectors classically affected by automation. A substantial displacement of labour will also challenge education and training institutions along with health, professional services, food and hospitality, and retail sectors if these prognoses on automation eventuate.

How are we to balance this predicted crisis of work against the capacities of technologies that drive current developments in automation specific to artificial intelligence (AI) and machine learning? A logic of substitution by which the machine stands in for the human worker, performing her task in more efficient and less costly ways, drives anxieties over labour redundancy. But questions about machines displacing jobs are not just economic matters. They also raise existential and ontological issues about the relation between life and mechanisms. At a time when computational processing has the possibility to generate "autonomous modes of automated epistemic production," we have to ask, with M. Beatrice Fazi (2018, 2), whether techniques like machine learning give rise to a radically "*alien thought*," quite distinct from that of the human mental worker. If we recognize the "possibility of novel thought" as "specific to what algorithmic automation is and does" (10), we cannot reduce the question of the relation between living labour and the machine to one of substitution or simulation. Automated technologies exert a force on the world – a force predicated on struggle, on social and technical relations that often enough come into conflict, or clash ontologically with the all-too-human predicament of labour. What, then, is the politics that emerges from the confrontation with these technologies? One can run to the state or embrace the private sector, or indeed recognize the co-mingling of these entities and move across the two. We prefer instead to grapple with the operations of power in pursuit of political analysis and critique that engages techniques and practices of world making.

Automation makes futures

We all know the science fiction scenario where robots take over. Perhaps this scenario does not eventuate. But what happens if automation takes over future planning? Automated planning is a branch of knowledge engineering that uses machine learning to synthesize computationally "ordered sets of actions that perform a given task" (Jiminéz et al. 2012, 433). Not restricted to the operations of robots and unmanned vehicles, these techniques serve up plural scenarios for strategic planning in corporate worlds. Promoted on LinkedIn – the purest of social networks, where making connections answers only the purpose of making connections – automated

scenario planning is a "future governance model" replete with "real-time ongoing reviews" and "metrics to drive pragmatic and informed decisions" (Coull 2016). As Stefano Harney quips in an interview with Michael Schapira and Jesse Montgomery (2017), "most managers have already been replaced by machines . . . We know they work not only within the parameters of an algorithm but with its predictions and prescriptions. They are only there to implement and call it leadership."

One influential report predicts that 47 per cent of all U.S. jobs are at risk of automation over the next twenty years, with up to 55 per cent of current jobs in metropolitan areas expected to disappear (Frey and Osborne 2013). Policy organizations report similar figures for Australia, with estimates of 40 per cent of jobs susceptible to automation processes (CEDA 2015). Importantly, Frey and Osborne estimate that approximately 30 per cent of jobs in the U.S. will be offshored over this same period. Such a dynamic suggests that automation is not exclusively about technology replacing human labour, but is rather a transformation accompanied by an ongoing complexification of the international division of labour in which some human labour forces perform work equivalent to the machine. Amazon's Mechanical Turk is perhaps the best-known example of this, where workers undertake routine data entry tasks for low wages (Irani 2015). Deutsche Bank chief executive John Cryan is particularly blunt in his prognosis of automation in the finance sector: "The truthful answer is we won't need as many people," he says. "In our banks we have people behaving like robots doing mechanical things, tomorrow we're going to have robots behaving like people" (quoted in Noonan 2017).

Cryan's comment brings full circle the dynamic we registered with respect to Marx's discussion of workers and automation in the "Fragment on Machines." Marx (1973, 693) worried that automation subsumes labour "under the total process of the machinery itself," rendering workers as mere linkages and thus robbing them of the human qualities of living labour. The transformation of living labour "into a mere living accessory" of machinery, Marx writes, "posits the absorption of the labour process in its material character as a mere moment of the realization process of capital" (693). Now the inverse of this process unfolds. Having created dehumanized workers, capital now replaces them with machines themselves, not because machine learning or other techniques of artificial intelligence share the human intelligence of living labour but because the latter has been reduced to such a level that it can now be artificially emulated, with the implication that we now live in a world more stupid than ever. Intelligence has fled the scene, or rather been recalibrated within the horizon of engineering and task-driven requirements of institution and economy.

No longer is the automation of institutional processes a practice restricted to the factory or bank. One could make similar attributions to any number of workplace settings, including those supposedly dedicated to "creative" or "intellectual" production. The university and higher education sector, for instance, are not immune from these transformations. Indeed, they have been part of the vanguard in developments that collapse the distinction between human and machine. University executives devise strategic policy that is indistinguishable from one university to the

next. Highly trained academics perform mind-numbing data entry tasks as a core component of their teaching workloads. Data analytics on student attention spans across course materials determined by template-driven learning objectives drive curriculum design. One can safely surmise that the cognitive autonomy of humans is becoming secondary to the vanilla dreams of machine-driven efficiencies.

As already mentioned, part of the discourse on the automation of labour is framed within debates on economic globalization and the offshoring of services. Automation also provides a precedent for exercises of reshoring. These involve the return of productive processes displaced from industrialized nations during the golden decades of offshoring. They require labour-eliminating technologies that can operate independently of the high labour costs in these countries. The paradigmatic example of such reshoring is the 2017 opening of an Adidas "speedfactory" in the German town of Ansbach. Equipped with "sewbots" and 3D printers that take instructions directly from design software, this facility not only allows rapid switching between the manufacture of different products but also drastically shortens the supply chain. Customers in European markets can access recently (and even custom) designed goods more quickly while Adidas cuts transportation costs. According to an International Labour Organization report (Chang, Ryanhart, and Huynh 2016), such developments threaten 90 per cent of garment and footwear jobs in Cambodia and Vietnam. These changes also explain why a return to eighteenth century economic technologies like the tariff, whatever its popular political appeal, is unlikely to alleviate problems of unemployment in the wealthy world.

The tech-sector has been a key driver in promoting the vision of a future without employment, where robots and intelligent systems deliver the utopian imaginary of leisure-driven life tasked with perpetual consumption. Reminiscent of earlier visions such as those of André Gorz (1980), these scenarios raise questions of sustainable business turnover and social reproduction – including the prospect of a universal basic income. Advocacy of the social wage is no longer restricted to postcapitalist hopefuls or those who agitate for the refusal of work. Championing the universal basic income has become *de rigueur* for tech companies and platform capitalists who worry that the "piece work" they dole out will be insufficient to sustain or reproduce the precarious labour forces upon which they depend.

A dystopian narrative is a perhaps more pervasive one in which the near future is underscored by enormous technological disruption. The masses are cast adrift in a world without work, where automation, as Sigfried Giedion (1948) claimed of mechanization, takes command. Under these conditions, the security of liberal democratic governance becomes a footnote in the history of machine-driven economies. As Isabell Lorey (2015) notes, governing through insecurity becomes the norm. Finding a path beyond the "new catastrophism" (Urry 2016, 33–53) that haunts such future scenarios requires attention to empirical conditions within data industries at the centre of digital automation.

A countervailing claim envisages the reabsorption of labour through other means. Questioning the accuracy of reports declaring the termination of employment for many, critics from fields such as science and technology studies consider automation

in a longer historical cycle that has more often resulted in the development of new forms of work (Wajcman 2017). Feminist scholars have noted the ways in which modern technological development liberated women from the bonded labour of domestic chores, enabling entry into new roles in workforces transformed by automation (Huws 2014). Where left-wing variations of accelerationism (Williams and Srnicek 2016) seek to reconcile machine-driven economies with the challenges of living within a capitalist world system, Nick Land's (2014) nihilistic vision of separate continental spaces for variously educated workforces hit by automation dramatizes anxiety around Ballardian and ultimately racialized imaginaries. It's no surprise that Land has become the poster boy for alt-right in recent times.

The prospect of an abolition of labour and arrival of cybernetic socialism has undergone undulations of enthusiasm among the left since the 1950s (for a trade press account of such scenarios, see Greenfield 2017, 192–193). More recently, a left variation of accelerationism pins post-work futures and the liberation of time to computational advances in automation and an intensification of capital accumulation. The elimination of work envisaged by critics such as Nick Srnicek and Alex Williams (2016) accompanies a repurposing of the means of production to invent postcapitalist futures. We are far less enthusiastic about such visions. Not because we don't share some of their sentiments or dream of a better world, but rather we see no rapid erosion of labour on global scales accompanying the latest wave of alarm surrounding automation.

To claim that automation makes futures is to refuse this alarm in ways that register the capacity for automation to make multiple futures. It is to reject the vision of accelerationism as much as countervailing claims that automation has no political consequence. In the face of the predictive power of data analytics, it is to reclaim the future in the name of uncertainty and surprise. It is to restore to politics the virtues of patience and expectation. Finally, it is to recognize the generative and destructive potentiality of automation asserted not only in opposition to the qualities of labour but also across myriad spaces in different and not necessarily consistent ways.

Automation needs data

The agglomeration of data is central to developments in AI and machine learning. Systems complexity is as good as the volume of data drawn on as agents in the design of computational parameters. But data are not simply given. To the extent that data have a referential dimension, their provenance lies with the technical devices and parametric designs from which they stem. Managed and amplified by software applications like mySQL and Hadoop, the neo-positivist fallacy assumed of data by academics, administrators, and executives alike posits the referent as external to the operation. The creation of datasets, let alone the smoothing out of frictions, contingencies, and differences between them, requires labour, even if this labour itself submits to automation. Fatigue is a continuum across the work of humans and machines. Performance measures index economic value calculated in relation to the optimization of efficiency. Data, like the human body, never

rest so much as signal their availability to engines of extraction. A latency defines moments of temporary suspension: from retrieval to task, the body of data learns the lesson of relevance. Consigned to storage and the vulnerability of the archive – this is what awaits both data and life deemed without purpose.

Critical data studies have taught us that data are never raw (Gitelman 2013). But how are data "cooked" within the circumstances of their collection, storage, and transmission? Such questions are vital for assessing the data politics that animate current automation. AI and machine learning use recursive techniques to update datasets in ways that allow them to evolve and improve their functionality. While statisticians and early data scientists worked with "sample" datasets, the evolution of so-called big data means that automation can now incorporate massive amounts of data in all their granularity, nuance, and detail. Big data analytics is all about the cut (Amoore and Piotukh 2015). Without partitions, there is no analysis. Context, or what we refer to as storage and the archive, jeopardizes the authority of decision. The algorithms that make decisions around the allocation of resources related to urban planning, for instance, need a constant supply of data. But not just any data will do. Data must be groomed for machine learning. There is a wide variety of techniques employed to make data ready for processing: selection, formatting, cleaning, scaling, decomposition, aggregation. Evelyn Ruppert (2017) refers to the deployment of these methods as the "crafting" of automation, emphasizing the role of metadata (or data about data) in establishing practices to secure trust in data. In recent times, the automation of such data preparation tasks through the establishment of data quality and governance rules has become a new front of extraction. Automation prompts more automation.

Xerox (2015) corporation invests in the continuity of craft and automation: "We believe that craft and automation are not opposites but complementary forces," declares a company publication. Here we see the ethos of the guild and the arts and craft movement celebrated by William Morris updated for twenty-first century corporate culture. But whatever efforts of skill and care are required for data integration exercises that resist the leverage introduced by metadata, the shaping of data is never merely a technical exercise. No dataset is neutral. Data absorb the social biases of the contexts in which they are generated and collected. The much commented-upon nexus of data and race is only the most obvious example of this. Algorithms are building new infrastructures of racism – for instance, in policing or credit rating. Consider the way financial institutions draw data from "racialized sources" such as local public records, social networking companies, academic records, mobile phone usage, non-financial payment histories, and even psychometric testing to construct credit ratings for the "unbanked" or those without credit records or files (Aitken 2017). Whether automated or human, the work of crafting data occurs within institutional settings and cultures that bind integration and analytics to economy and space.

Data analysis may turn nerds into celebrities, but the sociality of training and professionalisation of skills peculiar to the nineteenth century guild movement is nowhere to be found in the data extraction industries. As demonstrated by the

Cambridge Analytica–Facebook controversy of early 2018, the chief motive for corporations in the data economy is to decouple service from oversight by regulatory authorities. Once capital accumulation is unburdened of the subtractive imposition of state taxation regimes and restrictive legislation on data privacy, the economy and spatialization of data are free to play to the highest bidder. Politics becomes the art of the deal, not just the negotiation of distraction or the mobilization of preemption.

The launch of OpenAI in 2015 by Silicon Valley entrepreneurs to benefit "humanity as a whole" and guard against the "existential threat" of artificial intelligence is another exercise in cynical opportunism moonlighting as a data sharing exercise for all until the IPO, when the winner takes all. The question to ask of this organization is not about the ownership of the algorithms that run AI and machine learning platforms but the ownership of the datasets on which such platforms run. Data have become a kind of currency ("data is the new oil"). This is one reason why making data "open" or free to share has become a kind of evangelistic movement, crossing activist groups, software projects, and government institutions. As Nathaniel Tkacz (2012, 399) argues, such openness "actively works against the development of a political language – if we take the political to extend beyond questions of just governance to the circulation and distribution of power and force."

The deep imbrication of openness with ideals of political liberalism gives rise to hopes that making data publicly available will somehow generate citizens who use that data to inform action and debate. We are less hopeful that open data politics will produce such subjects. Nonetheless, a subject persists no matter the force of automation and AI, which will never be total as long as human life remains. This is a subject produced in and through struggle. At stake is a political subjectivity that holds no default allegiance to the liberal democratic state and its protocols of expression. The political subject we speak of here does not await authorisation to act. Nor does it celebrate its own right to have rights, to privacy for example. Instead, this subject seeks to understand how regimes of property, both public and private, limit the exercise of freedom. This subject does not shun technology. Rather, it subsists in a world of technical objects, collectively directing the machine ensemble in ways that contest and modulate the interiority of algorithmic decisions with the force of the outside. Data politics are not exclusive to the claiming of rights so much as the production of subjectivity within environments whose data architectures register conflicts between the politics of decentralization-centralization and the impossibility of pure distribution.

Aside from questions about what data are made open, how they are made available, and how they are used, the issue of how such openness leaves undisturbed (or even supports) proprietary regimes of data ownership and control needs to be broached. Opening data makes no challenge or difference to the speculative economies that surround the transfer of data from one party to another in corporate worlds. Indeed, European Commission research (Berends et al. 2017, 7) – itself outsourced to consulting firm Capgemini – emphasizes how private companies "transform open data from raw material into a service or product." This is

the business model of open source software companies. Little surprise that governments and intergovernmental organizations celebrate the passage of open data into private projects when it assists in reporting economic activity and growth. The requirement of open data to conform to standardized formats in order to be machine readable may establish what Felix Stalder considers "an important precondition for implementing the power of algorithms in a democratic manner" (2018, 167). We are more circumspect on this point. Certainly, questions of access and collective contribution to the production of a digital commons is facilitated by structures that support open data but, as Stalder goes on to note, examples such as the rise of the sharing economy demonstrate how labour enters a race to the bottom once open data become a commercial resource. In this sense, whoever sets the standards rules the world. The expansion of a data commons is not only a commercial precondition for data exploitation; technologies of automation also benefit from a deeper reservoir of data from which to calibrate their logics of operation (pattern recognition, preemption, prediction, and so forth).

Whatever openness or ownership protocols surround data, we need also to ask questions about the infrastructures that support that data. Key among these are data centres, whose capillaries of cables port data across racks of servers and cages of demarcation. Aside from a skeleton staff of technicians, security, and maintenance personnel, these are sites emptied of labour with labour distributed on the client end. While automation in the data centre requires data, there is not any great need for on-site workers. In this regard, the infrastructural object of the data centre offers one materialization and systemic instantiation of how automation scrambles and recasts regimes of labour. Not least it registers how data politics intersect with geopolitics, scattering labour forces connected at the client end of data infrastructures across political territories while also generating their own forms of territoriality and institutional power that parallel and rival the Westphalian architectures that haunt visions of statehood and internationalism.

Promoting the benefits of AI and machine learning, one recent tech-report notes: "There is also potential for AI technology to automate functions carried out by IT operations teams. Machine learning offers a way to manage infrastructure and react quickly to faults without human intervention" (Finnegan 2016). When human presence is marginal to the work of automation, the material form of the data society is not *ipso facto* one that has eradicated conditions of exploitation from the orbit of human existence and life in general. We all know how dependent platform economies are on the sociality of engagement. Whether it is the gang of FANG (Facebook, Amazon, Netflix, Google) coming out of Silicon Valley or the AI testing of facial recognition technologies for population management in China and internet services of Tencent and Alibaba, "participation" is the command action of exploitation required of humanity in the economy of data. Whether begrudgingly or with joy, participation is the primary technique of governing populations without purpose. Once agglomerated in datasets and made actionable by the instruction of code, the injunction to participate becomes secondary to the automation of operations. Data do not respond to interpellation.

Recognizing such participation as a front of exploitation is a first step to addressing and acting upon the discrepancy between living labour and the organization of social cooperation that is characteristic of the economic and social arrangements that spur capital valorization and accumulation in the contemporary world. Harnessing a critique of automated modes of production should not be something undertaken by an intellectual elite or vanguard. Indeed, the work of critique is just as much another passage of supply absorbed with indifference by the engines of data capitalism. Critique thus holds a recursive function that does nothing to transform the assumption of power, even if it retains a power to transform modes of cognition and political practice within the milieu of the social. This is the doubleness of critique: the production of value oblivious to the sign, and a signal of possibility and refusal to submit. The dependency of automation on data binds the machine world to the insubordination of labour even once automation learns the script of auto-generation.

Automation intensifies extraction

How then do data intercede between labour and capital? If from the viewpoint of media, automation presents the growing ubiquity of systems that appear autonomous or sovereign, from the perspective of capital it creates a scenario of increased dependence on so-called externalities. This predicament presents the paradox, or, better, relations of inverse proportionality, for which we must account in positioning labour with respect to automation. Automation is never bored, even if it is boring. Automation depletes the world of drama while it performs its routines of task work. Automation stages a vision of the world in which data is king, where the centre of gravity is the termination of labour. The seeming autonomy of automation instantiates the command of "sovereign media" in the realm of infrastructural territories (Adilkno 1998). At stake in such sovereignty is neither Hobbesian tacit consent nor Schmittian exception. Sovereign media are not about the projection of power as a search for truth, justice, or the capacity to coerce, but concern "only data which can be taken apart and reassembled in trillions of bytes" (Lovink and Richardson 2001). Sovereign media assume power through oblivion to the outside. This is also their weakness. Relations of dependency ruin the machine dream of perpetual accumulation. This is the revenge of labour.

In presenting capital with new frontiers of accumulation, data offer capital an outside to prospect and drawn upon. Here we can refer to Rosa Luxemburg's *The Accumulation of Capital* (2003), which argues that accumulation requires something more than a society composed of capitalists and workers. If, for Luxemburg, this outside comprised non-capitalist territories that colonialism could annex, the fronts offered by data are more profuse. Most of us are already familiar with the naming of these edges as economic sectors: finance, logistics, manufacturing, retail, security, healthcare, and urban "solutions."

The metaphor of data as a raw resource is ubiquitous enough, present whenever we hear talk of data mining. It is important to highlight the moment of extraction that this metaphor registers, even as we question that data can ever be raw and

plumb the ruptures as much as the continuities between resource extraction and data mining. As many have pointed out, what is extracted from the resource of data is not naturally given wealth but forms of social cooperation. Just think of the data extracted from social media use or the movement of pedestrians through a city. As in the cases of resource mining, monocultural agriculture, or other more traditionally extractive activities, such extraction also has its productive sides. Farming, milling, and modelling are metaphors often used to describe processes that turn such data into usable products for AI and machine learning. Nonetheless, we must also analyze the moment of extraction, since it is in this moment that the arrangements of data, labour, and capital specific to the current practices of automation become apparent.

Analyzing the situation in this way furnishes us conceptual tools with which to distinguish present frontiers of automation from those pertaining in the industrial (and national) moment of capitalism. In the factory, capital directly organized the materiality of productive cooperation and the introduction of the fixed capital of automation subsumed the variable capital of labour and subjected it to processes of intensification. Labour increasingly gravitated toward a state of equilibrium with the machine. Under present conditions of automation, capital is essentially indifferent to the relations of cooperation upon which it draws. Labour remains a condition of possibility for capital, but capital does not directly organize the relations of social cooperation upon which data economies rest. Consider the collection of data about users by social media companies, the financial hedging based on the slicing up of debts held by subprime mortgagees, or the ways in which the logistical coordination of supply chains push back costs to producers who can use whatever methods they like to keep prices low. In each of these instances, there is a moment of extraction enabled by the production and transfer of data. Clearly, modes of social organization are diverse across the different scenarios and practices that generate data for capital's engines. This diversity poses challenges for political organization, which in the current conjuncture necessarily grapples with the heterogenization of labour, across spatial parameters and legal differences as much as forms of insecurity and precarity that challenge the employment relations that prevailed in industrial modernity.

Automation amplifies abstraction. No longer can labour assume the possibility to enter a room to negotiate conditions of work and recompense for the toil of bodies and brains. There is no authority to arbitrate when automation glides on zero-hour contract jobs. No doubt, there are still unions with varying degrees of interventionist power – organized labour forces in the maritime, warehousing, and transport industries are among those most able to claim victories in disputes over conditions. But once data are captured, the automation of accumulation proceeds apace, absorbing and valorizing social knowledge and life in ways that extend way beyond delimited workplace relations. The classical repertoire of organization and protest no longer aligns with contemporary modes of capital accumulation. #Deletefacebook only generates data for Twitter. The boycott becomes just another means of capture. Even those Uber workers who use the app to organize makeshift stop-works or go-slows are bound to the data routines that make such assembly possible.

Sabotage seems an attractive possibility in these situations. But sabotage doesn't necessarily constitute a moment of organization adequate to new conditions, it just blocks the conduits for a while, spurring workarounds and securitizations on the part of capital. Likewise, the classical strike is not an effective weapon for labour against the distributed architectures and automated agents of data valorization. There is no there there, even if the infrastructures of capital are entirely concrete. The question of organizational form is key here. We must confront these conditions if we want to invent modes of labour organization that are adequate to automation and economies of data extraction.

Only by probing the ruptures and contradictions of capital within systems of data extraction can sites of struggle – protocological, social, legal, topological – manifest as conditions of possibility for new forms and techniques of organizing. A new conceptual language immanent to technical operations has the potential to align the politics of organization with the ongoing conflicts of extraction and exploitation. We need to recognize, for instance, that the fraction of capital that "unlocks value" from high frequency financial trading does not seek to organize the future productive activities upon which automated financial trades hedge. We also need to understand how productive processes in factories or transport hubs that are driven to the bottom by automated routines of logistical coordination. Similarly, the fraction of capital that automates using data generated by shoppers, patients, or social media users does not directly organize patterns of behaviour among these subjects. Rather, capital extracts a quota of value produced within these relations. Extraction, in this sense, names the forms and practices of valorization and exploitation that materialize when operations of capital encounter patterns of human cooperation and sociality external to them.

If data centres are the automated mills that enable such extraction, we have to ask questions about the displacement of the human from this machinery of capital. One way to characterize the social cooperation from which capital creates value by extracting data is to say it involves the "production of humans by humans" as opposed to the production of commodities in the industrial factory (Marazzi 2005). The production of subjectivity and forms of life are increasingly central to capitalist valorization. If we introduce automation to this scenario, it becomes tempting to view the frontiers of such valorization as producing old battle lines between the human and machine: humans cooperate to make value and machines extract that value as data.

We would do better to heed Félix Guattari's (2011) notion of the machinic assemblage, which brings humans and nonhumans into dynamic interaction on a single ontological plane. In the present wave of automation, media may assume a sovereign prerogative, but only because they are shackled to their environments. This is the crucial difference between automation and mechanization. We know this from cybernetics. Automated technologies are not machines that do the same thing every time. Like kinetic sculptures, the making of which provides one of the most dynamic fields for the critique of current automation, they operate by

interacting with their surroundings. We must further explore the nexus of media and environment to position data infrastructures in their geopolitical as much as their natural settings.

Automation adapts to environments

The informatic machines that drive today's automation are different creatures from the thermodynamic machines that populated the nineteenth century factory. Sure, digital technologies consume electricity and other kinds of power – in vast amounts when aggregated or observed in the economies of scale created by data centres. These machines also produce their own forms of waste, including heat, noise, and carbon dioxide. This energy consumption and waste production can never be forgotten when considering the relation of computer-based automation to the environment, as numerous practitioners of green media studies (Cubitt 2017; Miller and Maxwell 2012) remind us. Yet digital technologies run on the fantasy that they leave no trace in the atmosphere. As Alexander Galloway (2018) puts it, "the essence of the informatic machine is found in form, not energy or presence." From this perspective, the precedent for digitally based automation lies not in the spinning jenny as much as the Jacquard loom, which introduced patterns into mechanized weaving through information printed on cardboard ribbons – a version of the punch cards used in computing well into the 1970s. The computer purports to transcend its body and environment, to occupy a realm of abstraction that negotiates relations of exchange, differentiation, and equivalence. How do we understand this claim to autonomy, the computer's desire to be free of its environment?

No matter how urgent it may be in the age of the Anthropocene, moralizing about energy and waste can only take us part way towards answering this question. Automated technologies undeniably interact with their environments in thermodynamic ways. They also rely on their environments as sources of information, as the very basis of their so-called autonomy. Automation is not a mechanized routine that performs the same action over and again. Contrary to visions that paint current transformations as updated versions of some old industrial revolution, we no longer live in Charlie Chaplin's *Modern Times*. Rather, automation lives by adapting to its environments, even if only in input-output ways mediated by patterns of machine learning.

Media as environment define societies of automation. As German media philosopher Erich Hörl (2017, 9) maintains, environmentality is the emergent mode of governmentality in which media are capable of self-organization. Literary theorist and historian of cybernetics N. Katherine Hayles (2017) extends this baseline of computational control in her study of the "cognitive nonconscious" to examine automated traffic systems in Los Angeles where algorithms, processors, and databases extract patterns and modify operations in response to changing externalities. Such theoretical insights bear on a study of how automated technologies govern or displace labour within manufacturing and service industries.

We also find a key signal here in the vulnerability of environmentality as an emergent mode of governance. The backup of data across facilities clustered in a region or dispersed across a global geography of infrastructure tells us not just that duplication makes data secure, but also that failure is a component part of designing environments receptive to contingency. In seeking to close the disruptive force of externalities, data centres embody the core concept of "noise" trawled over in the renowned Macy conferences on cybernetics running from after World War II and finishing in 1960. Organized as informal conversations and published in a series of proceedings, an extraordinary range of figures representing the cutting-edge of research in disciplines including mathematics, biology, architecture, anthropology, medicine, communications, psychiatry, and linguistics attended the Macy conferences. Reading across the proceedings of these years (Pias 2003), it is striking how preoccupied the debates were with the constitutive role of noise or feedback as both a reflexive and recursive operation within cybernetic systems. While never articulated as such, we can now understand this fascination as an emergent recognition of environment as a complex of relations that govern and make operative communication and cognition, reproduction and dissemination.

Understanding the environmentality of automation in this way gives us a new angle on the relation of labour to digital technologies. Usually labour seems far too human-centred a category to populate arguments about the urgencies of environmental change. Insofar as automating technologies are nonhuman agents, they share something with other nonhuman actors such as animals, plants, and rocks. Not surprisingly, there is no shortage of claims that automation benefits the environment, either by making production leaner (a claim that forgets the energy burnt and waste emitted by data infrastructures) or through the automation of environmental control – energy monitoring systems and the like. Machines and nature ally against the productive vanities of human labour. But if we understand environment as the noise of automation, we situate it, like labour, as a constitutive outside of automating systems. Far from alienating labour from the world of nature and thus inverting the Lockean schema by which labour turns nature into property, we need to consider how labour fits into the environment in ways that do not reduce the latter to a resource for exploitation. Environment is not only nature. For this reason, we find Edward Burtynsky's (2006) aestheticization of manufactured landscapes more compelling testaments to the challenges of environmentality than celebrations of actors and networks that fix agency on objects such as mushrooms that might fend off the end of the world (Tsing 2015). Automation needs a politics that questions survivalism. The panoramic views of human-altered landscapes that sweep across Burtynsky's film are just one way of registering the aesthetics of automation.

Automation fails

"Fuck content," writes Michael Rock (2013) in an eponymous essay that celebrates the designer's purview to shape. We might as well take this as a mantra

for automated aesthetics. Just think of the personalized social media feed, an automated object arranging updates and info, viewpoints and videos. What counts is the form. Whether the content is hate speech, fake news, schlock, or banality, the feed serves it up just the same, at least until other automated systems detecting porn and other vices "recognize objectionable parts by seeing a sizable mass of them in order to infer their relations" (Steyerl 2017, 36). Once the assembly of masses on the street prompted anxiety on the part of authorities. Nowadays spikes in data draw the attention of algorithms in search of irregularities. We also see the triumph of form in Snapchat's face filters, so enamoured by bored teenagers. It's all about how the doggy ears, flower crown, or rainbow vomit automatically adjust to the selfie layered below. The Viola-Jones facial recognition algorithm is the real star, not the adolescent indulging the narcissism of facial metamorphosis. Machine vision formalizes the world as an algorithmic topology from which a variety of contents – trackable objects, readable numbers, or parseable text – can be extracted. As Jacques Rancière (2004) reminds us, aesthetics is never purely about beauty or style, but inherently political in that they determine who gets to speak and who remains silent, what gets visualized and what remains invisible. Beyond the logic of representation is the politics of operation.

Yet the politics of operation bring with them an aesthetics of failure. There is a failure of aesthetics to register the operationality that so often glides beneath the threshold of sensation and perception (*aesthesis*). Operations incorporate failure as a generative force. In this respect, operations never end even if their component parts, elements, and dynamics transmogrify and adapt according to the theatre of failure. Faults and failures are a regular feature of operations, so routing mechanisms must make use of redundancies offered by hardware. The typologies of fault tolerance move us into different classifications of failure: failure type, failure region, failure neighbourhood, failure mode, and failure time (Yang Liu et al. 2013, 51–53). Failure in the form of "crashes and viruses, bloatware, malware, and vaporware," as media theorist Florian Cramer (2005, 9) notes, comprises the "irrationality of rational systems." Once failure becomes normative within the spectrum of operations, we might lament the way in which the machine smooths over or absorbs contingency. To harbour such nostalgia is to overlook the persistence of labour. System failure provides labour with a shadow in which to cloak itself from regimes of inspection. Disruption to the storage and processing of data includes power outages, cooling system failure, the manufacturing of defective switches and faulty devices, security glitches, memory errors, zombie servers, and the like.

Despite the trend toward automating data infrastructures, the constitutive outside of labour remains as an externality not yet entirely congruent with the calibrated logic of automation. No matter that micro-contracting platforms corrode the distinction between automating tasks and the font of human labour, ingenuity persists in the human capacity to beaver out the root cause of a problem in the event of machine failure. Not limited by parameters that struggle with unforeseen contingency, the human retains what Hayles (2017) terms "the power

of the cognitive nonconscious." In pursuit of technical agency on par with neural networks of the human brain, research and development in AI and machine learning has still to conquer this last frontier that makes human labour a viable proposition in regimes of capital accumulation. Hayles: "the gap between biological nonconscious cognition and technical nonconscious cognition still yawns as wide as the Grand Canyon on a sunlit morning" (3).

Critical ethnographies of infrastructure (Star 1999) teach us that the material substrates that underlie and enable contemporary life become evident only when they break down. Brian Larkin (2013) questions this perspective by pointing to the aesthetics of infrastructure, its monumental or even fetish-like qualities. We are all familiar with the banal failures that automation wreaks on our daily lives – for example, automated software updates make our image files and websites crash. Automation can also become a fetish that operates beyond its technical functions. It can stand for capital's fantasy of doing away with labour and routing around the subject. Couched in the idiom of object-oriented ontology or coded in object-oriented programming languages like Python, we must balance this fantasy with the knowledge that living labour remains a constitutive outside of automation.

Automation also fails in this sense. At some point, it has to measure up against the human capacity to refuse. Such refusal can manifest itself in acts of infrastructural sabotage, which take the form of cyber espionage, election rigging, hacking into systems, and so forth. It is also manifest in the world of data centres, where we find an industry of infiltration and misappropriation not advertised in job descriptions on bulletin boards. The culture of darknets is replete with trade secrets, high-level organizational data, backend advice on accessing secure systems, etc. Sites like 4chan.org are sufficiently emboldened to make enough of this sort of material openly public (Coleman 2014). While in the world of finance, the phenomenon of so-called dark pools generates a "bordered space of private financial transactions that is increasingly free from national and international regulatory authorities" (Sassen 2013). A more political yearning for autonomy is evident in the repurposing of data infrastructures in ways that tread a delicate line between the conviction that human design is sovereign and the realization that machinic automation increasingly takes command.

The automation of society and economy registers the shift from the governance of supply chains and labour regimes subject to real-time key performance indicators measured within enterprise resource planning software to higher-level automated systems designed around the cybernetic principle of noise and feedback of externalities. With logistical media, the programmer and engineer effectively occupy the seat of control. They design the parametric architectures and code the machines. The advent of AI and machine learning marks the arrival of autonomous media – computational systems independent of human oversight and able to exert a form of sovereign authority over the organization and management of society.

Ultimately, the question of how labour transitions to a society of automation is a political question, and not only because the subject of labour was for a century or more imagined as a political actor that could transform the social. We must

also recognize how capital in its own right has emerged as a political actor on the global stage, producing territories, legislations, and fronts of violence. Indeed, conjuring a society irreducible to regimes of measure is a collective undertaking that requires a knowledge of operations in a world that already seems to have complied with all the terms and conditions. Such an undertaking can never see automation as a moment of total closure. The central paradox by which automation reduces labour to a mere abstraction but makes the machine a living entity remains in play. Automation fails to close this opening. Recognizing the machine's capacity to become a subject is no less crucial than registering the human subjectivity that capital attempts to route around. Automation defines a field of political struggle. This is why AI and machine learning are unlikely to resolve the time-honoured question of whether technology eliminates employment or moves it to other realms. Automation's magic will continue to serve as a proxy for its internal clockwork. The politics of reorientation must script new horizons as externalities not yet trafficked through global infrastructures of control.

References

Adilkno. 1998. *The Media Archive: World Edition*. New York: Autonomedia.

Aitken, Robert. 2017. "'All Data is Credit Data': Constituting the Unbanked." *Competition & Change* 21 (4): 274–300.

Amoore, Louise and Volha Piotukh. 2015. "Life beyond Big Data: Governing with Little Analytics." *Economy & Society* 44 (3): 341–366.

Benjamin, Walter. 1969. "Theses on the Philosophy of History." In *Illuminations*, edited by Hannah Arendt, translated by Harry Zohn, 253–264. New York: Schocken Books.

Berends, Jorn, Wendy Carrara, Wander Engbers and Helen Vollers. 2017. *Re-Using Open Data: A Study on Companies Transforming Open Data into Economic & Societal Value*. European Data Portal Project. Brussels: European Commission.

Braverman, Harry. 1974. *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century*. London: Monthly Review Press.

Burtynsky, Edward. 2006. *Manufactured Landscapes*. Zeitgeist Films.

Chang, Jae-Hee, Gary Rynhart and Phu Hunyh. 2016. *ASEAN in Transformation: The Future of Jobs at Risk of Automation*. International Labour Organization, 1 July. www.ilo.org/wcmsp5/groups/public/---ed_dialogue/---act_emp/documents/publication/wcms_579554.pdf.

Coleman, Gabriella. 2014. *Hacker, Hoaxer, Whistleblower, Spy: The Many Faces of Anonymous*. London: Verso.

Committee for Economic Development of Australia (CEDA). 2015. *Australia's Future Workforce*, June. www.ceda.com.au/Research-and-policy/All-CEDA-research/Research-catalogue/Australia-s-future-workforce.

Coull, Terry. 2016. "Today's CIOs Focus on the Right Things, but is the Right Attention Given to Exploring Future Opportunities?" *LinkedIn*, 18 October. www.linkedin.com/pulse/today-cios-focus-right-things-attention-given-exploring-terry-coull.

Cramer, Florian. 2005. *Words Made Flesh: Code, Culture, Imagination*. Rotterdam: Piet Zwart Institute.

Cubitt, Sean. 2017. *Finite Media: Environmental Implications of Digital Technologies*. Durham: Duke University Press.

Fazi, M. Beatrice. 2018. "Can a Machine Think (Anything New)? Automation Beyond Simulation." *AI & Society*, 12 February. Advance online publication. doi.org/10.1007/s00146-018-0821-0.

Finnegan, Matthew. 2016. "Why machine learning could be the next frontier for data centre operations," *Computerworld*, 24 October. www.computerworlduk.com/infrastructure/machine-learning-is-next-frontier-for-data-centre-operations-3648124/.

Frey, Carl Benedikt and Michael A. Osborne. 2013. "The Future of Employment: How Susceptible are Jobs to Computerisation?" University of Oxford, 17 September. www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf.

Galloway, Alexander. 2018. "Anticomputer." *Culture and Communication*, 19 March. http://cultureandcommunication.org/galloway/anti-computer.

Giedion, Sigfried. 1948. *Mechanization Take Command: A Contribution to Anonymous History*. New York: Oxford University Press.

Gitelman, Lisa, ed. 2013. *"Raw Data" is an Oxymoron*. Cambridge: MIT Press.

Gorz, André. 1980. *Abschied vom Proletariat*. Frankfurt: Europäische Verlagsanstalt.

Greenfield, Adam. 2017. *Radical Technologies: The Design of Everyday Life*. London: Verso.

Guattari, Félix. 2011. *The Machinic Unconscious: Essays in Schizoanalysis*. Translated by Taylor Adkins. Los Angeles: Semiotext(e).

Hayles, N. Katherine. 2017. *Unthought: The Power of the Cognitive Nonconscious*. Chicago: University of Chicago Press.

Hörl, Erich. 2017. "Introduction to General Ecology." In *General Ecology: A New Ecological Paradigm*, edited by Erich Hörl with James Burton, 1–73. London: Bloomsbury Academic.

Hoskins, Tansy. 2016. "Robot Factories Could Threaten Jobs of Millions of Garment Workers." *The Guardian*, 16 July. www.theguardian.com/sustainable-business/2016/jul/16/robot-factories-threaten-jobs-millions-garment-workers-south-east-asia-women.

Huws, Ursula. 2014. *Labor in the Global Digital Economy: The Cybertariat Comes of Age*. New York: Monthly Review Press.

Irani, Lilly. 2015. "Difference and Dependence among Digital Workers: The Case of Amazon Mechanical Turk." *South Atlantic Quarterly* 114 (1): 225–234.

Jiminéz, Sergio, Tomás De La Rosa, Susana Fernández and Fernando Susana Fernández. 2012. "A Review of Machine Learning for Automated Planning." *Knowledge Engineering Review* 27 (4): 433–467.

Land, Nick. 2014. "Teleoplexy: Notes on Acceleration." In *#Accelerate: The Accelerationist Reader*, edited by Robin Mackay and Armen Avession, 509–520. Falmouth: Urbanomic.

Larkin, Brian. 2013. "The Politics and Poetics of Infrastructure." *Annual Review of Anthropology* 42 (1): 327–343.

Leontief, Wasily and Faye Duchin. 1986. *The Future Impact of Automation on Workers*. New York: Oxford University Press.

Lorey, Isabell. 2015. *State of Insecurity: Government of the Precarious*. Translated by Aileen Dereig. London: Verso.

Lovink, Geert and Joanne Richardson. 2001. "Notes on Sovereign Media." *Subsol*. http://subsol.c3.hu/subsol_2/contributors0/lovink-richardsontext.html.

Luxemburg, Rosa. 2003. *The Accumulation of Capital*. Translated by Agnes Schwarzschild. London: Routledge.

Malthus, Thomas. 1836. *Principles of Political Economy*. London: W. Pickering.

Marazzi, Christian. 2005. "Capitalismo digitale e modello antropogentico del lavoro." In *Reinventare il lavoro*, edited by Jean-Louis Lavalle, Christian Marazzi, Michele La Rosa, and Federico Chicchi, 107–148. Roma: Sapere 2000.

Marx, Karl. 1973. *Grundrisse*. Translated by Martin Nicolaus. London: Penguin.
Marx, Karl. 1977. *Capital*. Translated by Ben Fowkes. New York: Vintage.
Meade, James E. 1964. *Efficiency, Equality and the Ownership of Property*. London: George Allen & Unwin.
Méda, Dominique. 2016. *The Future of Work: The Meaning and Value of Work in Europe*. International Labour Organization Research Paper, no. 18. October.
Miller, Toby and Richard Maxwell. 2012. *Greening the Media*. New York: Oxford University Press.
Noonan, Laura. 2017. "Deutsche Boss Calls for 'Revolutionary Spirit' as Robots Take Jobs," *Financial Times*, 6 September. www.ft.com/content/398836c4-92e0-11e7-a9e6-11d2f0ebb7f0.
Pias, Claus, ed. 2003. Cybernetic-Kybernetik: The Macy Conferences, 1946–1953, Vol. 1. Zürich-Berlin: Diaphanes.
Rancière, Jacques. 2004. *The Politics of Aesthetics: The Distribution of the Sensible*, translated by Gabriel Rockhill. London: Continuum.
Regalado, Antonio. 2012. "When Machines Do Your Job." *MIT Technology Review*, 11 July. www.technologyreview.com/s/428429/when-machines-do-your-job/.
Ricardo, David. 1951. "Principles of Political Economy and Taxation." In *The Works and Correspondence of David Ricardo*, edited by Piero Sraffa. Cambridge: Cambridge University Press.
Rock, Michael. 2013. "Fuck Content." In *Multiple Signatures: On Designers, Authors, Readers and Users*, edited by Michael Rock, 45–56. New York: Rizzoli International.
Ruppert, Evelyn. 2017. "The Crafting of Automation." RMIT University Digital Ethnography Research Centre, 29 November. http://digital-ethnography.com/events/crafting-automation-evelyn-ruppert-public-lecture/.
Sassen, Saskia. 2013. "When Territory Deborders Territoriality." *Territory, Politics, Governance* 1 (1): 21–45.
Say, Jean-Baptiste. 1803. *Traité d'Économie Politique*. Paris: Chez Deterville.
Schapira, Michael and Jesse Montgomery. 2017. "Stefano Harney (Part 1)." *Full Stop Quarterly*, 8 August. www.full-stop.net/2017/08/08/interviews/michael-schapira-and-jesse-montgomery/stefano-harney-part-1/.
Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: William Strahan and Thomas Caldwell.
Star, Susan Leigh. 1999. "The Ethnography of Infrastructure." *American Behavioral Scientist* 43 (3): 377–391.
Stalder, Felix. 2018. *The Digital Condition*. Translated by Valentine A. Pakis. Cambridge: Polity Press.
Steyerl, Hito. 2017. *Duty Free Art: Art in the Age of Planetary Civil War*. London: Verso.
Tkacz, Nathaniel. 2012. "From Open Source to Open Government: A Critique of Open Politics." *ephemera: theory & politics in organization* 12 (4): 386–405.
Trott, Ben. 2017. "The 'Fragment on Machines' as Science Fiction; or, Reading the *Grundrisse* Politically." *Cambridge Journal of Economics*. Advance online publication. doi:10.1093/cje/bex079.
Truitt, Elly R. 2015. *Medieval Robots: Mechanism, Magic, Nature, and Art*. Philadelphia: University of Pennsylvania Press.
Tsing, Anna. 2015. *The Mushroom at the End of the World: On the Possibility of Life in Capitalist Ruins*. Princeton: Princeton University Press.
Urry, John. 2016. *What is the Future?* Cambridge: Polity.
Williams, Alex and Nick Srnicek. 2016. *Inventing the Future: Postcapitalism and a World without Work*. London: Verso.

Voskuhl, Adelheid. 2013. *Androids in the Enlightenment: Mechanics, Artisans, and Cultures of the Self*. Chicago: University of Chicago Press.

Wajcman, Judy. 2017. "Automation: Is it Really Different this Time?" *The British Journal of Sociology* 68 (1): 119–127.

Xerox. 2015. *The Automation of Craft, the Craft of Automation*. Accessed 1 April 2018. www.xerox.com/downloads/services/ebook/printing_craft_to_automation.pdf.

Yang Liu, Jogesh K. Muppala, Malathi Veeraraghavan, Dong Lin, and Mounir Hamdi. 2013. *Data Center Networks: Topologies, Architectures, and Fault-Tolerance Characteristics*. Cham: Springer.

11

DATA'S EMPIRE

Postcolonial data politics

Engin Isin and Evelyn Ruppert

Postcolonial data politics

This chapter addresses a question that is rarely, if at all, raised about the ways in which data politics plays out differently in colonial, postcolonial, and imperial states and their respective force in international relations (Burbank and Cooper 2010, Gilroy 2004, Kumar 2017, Muldoon 1999). The chapter makes a case for a distinct "postcolonial data politics" to draw attention to how data politics plays out differently in the Global South than the Global North. We develop this case by first examining its conditions of possibility: the colonial power and knowledge (institutions, disciplines, objects, and subjects) that constituted and continue to shape postcolonial states and their relationships with imperial states. One such condition that we exemplify is how the quest for a British imperial census in the nineteenth century and its technologies of colonial government of counting, categorising, and ordering were inherited, reshaped, and reused by postcolonial governments. Understanding the constitutive force of this genealogy is key to interpreting how the vast amounts of data collected through the internet and devices continues yet reconfigures colonial logics and objects of knowledge.

In the second edition of *Imagined Communities* (1991) Benedict Anderson concludes that he was "hasty" and "superficial" in the original edition (1983) in assuming that twentieth-century postcolonial states were modelled after modern nineteenth-century European states (2006, 163). He suggests that to understand nationalism in postcolonial states, its genealogy should be traced to colonial governments instituted by imperial powers before the nineteenth century. He thinks that it is necessary to understand "imaginings of the colonial state". He admits that "this conclusion may seem surprising, since colonial states were typically anti-nationalist, and often violently so. But if one looks beneath colonial ideologies and policies to the grammar in which, from the mid nineteenth century, they were deployed, the lineage becomes decidedly more clear" (2006, 163). For students

of colonialism and imperialism, neither the proposition that postcolonial states inherited practices from colonial institutions nor that imperial states transplanted practices from colonial governments – something that Michel Foucault (2003, 103) had identified in 1975 as a "return" effect – would come as a surprise.

Anderson argued that this genealogy was expressed and deployed most prominently in three colonial institutions – the census, the map, and the museum: "together", he says, "they profoundly shaped the way in which the colonial state imagined its dominion – the nature of the human beings it ruled, the geography of its domain, and the legitimacy of its ancestry" (2006, 163). His analysis of each institution in imperial states that colonised Southeast Asia have been widely discussed as well as his claim that its lessons should have comparative value as it includes territories colonised by the all "white" imperial powers of Britain, France, Spain, Portugal, The Netherlands, and the United States (Appadurai 1993; Christopher 2009, 2008; Cordell, Ittmann, and Maddox 2010).

Anderson's analyses of the census (population), the map (territory), and the museum (memory) are pertinent for the argument we want to make in this chapter. As modes of knowledge, especially the census and the map, were not merely descriptive exercises that represented populations and territories but were performative technologies that literally produced them. As James Scott says, maps

> were . . . not just maps. Rather, they were maps that, when allied with state power, would enable much of the reality they depicted to be remade. Thus a state cadastral map created to designate taxable property-holders does not merely describe a system of land tenure; it creates such a system through its ability to give its categories the force of law.
>
> *(J.C. Scott 1999, 3)*

If the map was not merely a representation of a given territory, it came to constitute territory as an object of power: possession and dominion were synonyms of the colony. The same can be said about both the census and the museum. They were not merely representations of "population" and "memory" but practices through which they became objects of power. This was of course compellingly developed by Bernard Cohn (1996) in his studies of British imperial government in India. On the census Anderson largely focused on the production of ethnic and racial categories and how they helped shape the imagination of the nation constituted by the very categories imperial powers instituted.

It should be said that it is not only that the will to power and knowledge mobilised the census but also the constitutive performative force of that which it produced: population. The enormous amount of data collected, collated, interpreted, analysed, and disseminated about the colonies provided the ways in which the dominions and possessions were imagined in the sense Anderson always maintained: as produced. This is the performative sense in which the data produced about an object at the same time exceeds its will to power and attains constitutive powers in shaping and forming that object. This is certainly the sense in which Edward Said's (2003) critique of

orientalism as a discourse – practical, academic, and literary – constituting the orient specially but the colony generally highlighted. As Young puts it

> Said's use of the notion of a discourse to demonstrate the way in which forms of knowledge were constructed within a particular kind of language, which in turn was replete with all sorts of cultural assumptions, enabled Orientalism, and colonialism more generally, to be analysed as an ideological production across different kinds of texts produced historically from a wide range of different institutions, disciplines and geographical areas.
>
> *(Young 2016, 385)*

The most important lesson we have learned from political sociology and anthropology of empires is that while an empire that embodies a will to power may come to pass and its mode of dominion or possession may become postcolonial, the constitutive powers of knowledge (institutions, disciplines, objects, and subjects) continue to shape postcolonial states and their relationships with imperial states. Conversely, modern European empires never developed knowledge-power practices in isolation from their metropoles. Thus, we use the terms colonial, postcolonial, and imperial states to highlight their genealogy and insist on using "empire" neither as ubiquitous and omnipresent nor as a geographically contiguous and historically homogenous form of rule (Kumar 2010). Rather, we use "empire" to signify a form of rule whose performativity constitutes power relations between dominant and dominated institutions and as such can be geographically dispersed and historically heterogeneous (Bourdieu and Wacquant 1999).

Our argument is that the continuing and constitutive powers of knowledge (of population, of territory, and of memory) should have significant bearing on how we now think about the vast amounts of data collected through the internet and devices that are said to usher in a new era of data politics. We argue in this chapter that postcolonial data politics should be a distinct domain of analysis by focusing on how colonial dominions and possessions are now being reconfigured as objects of knowledge. We develop this argument in three stages. In the following section, we focus especially on the British Empire and its massive efforts to establish an imperial census beginning from the 1840s to the 1940s. This is to illustrate how the census represented a mode of data politics that produced colonial populations as objects of power. Then we will focus on two cases from contemporary data politics where postcolonial states are being increasingly brought under the orbit of massive data collection, collation, and interpretation regimes by new kinds of authorities whose mode may not be imperial yet whose form of rule distinctly is. Then we will conclude with some thoughts on resistances to data's empire and possibilities of decolonising data politics.

Governing peoples: biopolitics and empire

There has been a lively debate on empires over the last twenty years. There is no doubt that this debate owes a great deal to Said's *Orientalism* (2003) and *Culture*

and Imperialism (1994) – two books that shaped and framed the subsequent field of postcolonial studies (Young 2016). Perhaps counterintuitively, postcolonial studies opened up, amongst other things, the possibility of understanding the continuity between colonial and postcolonial states and the role of imperial states (those states that established conquered or settler colonies and dominions). What Anderson called his oversight – that he did not originally see the relation between colonial and postcolonial states, regarding the latter's nationalism as the negation of the former – is indeed an insight that arises strongly from postcolonial studies. It has been now widely debated that rather than "disappearing" especially, European empires have taken on new forms. Jane Burbank and Frederick Cooper (2010) have, for example, entirely shifted the ground by comparatively investigating empires as ongoing forms of rule for governing peoples and populations. Similarly, Krishnan Kumar (2017) focused on five world empires and their continuing presence in the contemporary world. To be sure, the conclusion to draw from this burgeoning debate and the shifting ground is not that "empires are alive and well" but that empire is a changing form of rule that shapes global population management and creates evolving forms of subject peoples (Hevia 2012, Ittmann, Cordell, and Maddox 2010, Pagden 2001, Steinmetz 2013). Within postcolonial studies the terms metropole and postcolony or "the Global North" and "the Global South" are used to indicate a deterritorialised geography where "Souths in the geographic North and Norths in the geographic South" are entangled (Mahler 2018, 19, Mbembe 2001). It is in this deterritorialised sense that we adopt the terms metropole, colony and postcolony when referring to empire as a form of rule.

There is so much more to say than we can in this chapter about this debate but we will briefly draw out its significance for our argument for a postcolonial data politics. We want to illustrate this by first returning to Michel Foucault who not only influenced figures such as Edward Said and postcolonial studies but also inaugurated studying "population" as an object of modern government – understood as a broad concern with the administration of things and people. Yet, as Anne Stoler (1995) famously argued, with the exception of a brief note on the "return effect" that we have mentioned earlier, Foucault did not concern himself with colonial government let alone colonial technologies of power such as the census, the map, or the museum in the colony (D. Scott 1995). One would have expected that studies on "colonial government" and especially the colonial census and the production of colonial populations would have flourished, but this happened only to a limited extent and there are not many studies of colonial populations and their principles of production especially in the context of imperial government (Christopher 2008, Cordell, Ittmann, and Maddox 2010, Ittmann, Cordell, and Maddox 2010). Early studies have not been followed through with detailed investigations especially outside India (Appadurai 1993, Cohn 1996, Ludden 1993, Kalpagam 2000b). It is well worth then revisiting Foucault (albeit briefly) on population, draw out the relations between biopolitics and "data politics" as we see it, and provide an overview of the British Empire's attempts at creating an imperial census as a prologue to postcolonial data politics.

In his Collège de France lectures on Society Must be Defended (1975–1976) and Population, Security, Territory (1977–1978), Michel Foucault (2007, 2003) outlines what he sees as the specificity of modern government. The publication of these lectures (in English) nearly twenty years after their delivery has been a revelation for those who took the last chapter of his *History of Sexuality* (1978) as a ground-breaking attempt to identify the specificity of modern government. In these lectures he more clearly outlines how he thinks the concept of government acquires a broad meaning in the sixteenth century when the verb "to govern" functions in a wide range of domains to indicate any benevolent or prescriptive activity to command the movements and subsistence of people (2007, 122). Foucault concludes that "one thing clearly emerges through all these meanings, which is that one never governs a state, a territory, or a political structure. Those whom one governs are people, individuals, or groups" (2007, 122). The target and object of government is always a people, individuals, or groups. Yet, at this point, in the sixteenth and seventeenth centuries, governing peoples is a sovereign exercise: it is direct, violent, and unforgiving.

Foucault argues that by the second half of the eighteenth century a new form of power adds another meaning of government, which he eventually calls "disciplinary". He does not think that "sovereign" and "disciplinary" mechanisms of power are to be juxtaposed against each other. Rather, these two mechanisms of power work with different rationalities. Still, this is probably not the most original argument of Foucault. It is when he becomes aware that toward the end of the eighteenth and early nineteenth centuries a new mechanism of power emerges. In a much-quoted statement Foucault says "unlike discipline, which is addressed to bodies, the new non-disciplinary power is applied not to human-as-body but to the living human, to human-as-living-being; ultimately, if you like, to human-as-species" (Foucault 2003, 242). If disciplinary power is concerned with the function, movement, and fitness of the body, regulatory power is concerned with the birth, death, and health of the species-body. Foucault calls this emerging regime of regulatory mechanisms of power that are concerned with the species-body as biopolitics – so-called because of its concern with natality, mortality, and fertility (Foucault 2003, 243). If then the body is the problem of disciplines, then population becomes the problem of regulation (Foucault 2003, 245). The key argument Foucault makes is that although population was an object of power and knowledge before the late-eighteenth century, it became an object of management that required new techniques of data collection (e.g., census) and new methods of analysis (e.g., demography, statistics). Although Foucault is not always consistent, we do not think that he offers these three logics of government – sovereignty, discipline, and regulation – as supplanting or displacing each other but that the specificity of modern government consists in their multiple and intersecting deployments appropriate to each target of government.

Of course, much has been debated over Foucault's lectures especially on territory, population, and security over the last decade or so and we do not aim to discuss the main issues of agreement and disagreement. From our perspective we

want to note that despite the substantial debate over Foucault's claims about early modern and modern European states when European empires were also accumulating territories, creating populations and advancing competing claims over (and warring for) sovereignties in the colonies, his claims remain crucial for understanding the development of colonial government (Curtis 2004, Kalpagam 2000a, 2001, D. Scott 1995, Wilson 2011). What Ian Hacking (2015) identified as the avalanche of printed numbers between the 1820s and 1840s when "population" was invented is precisely the period in which British imperial government identified its colonial populations differently. We want to illustrate this briefly with the efforts to conduct an imperial census during the British Empire.

Consistent with Foucault's claim that population becomes a new kind of object of government in the late-eighteenth and early-nineteenth centuries, the British Empire had a renewed interest in its colonial dominions and possessions and news ways of accounting for them. As Foucault would also emphasise, this was not the first time that states had an interest in their populations but their modes of accumulating data about them signalled something different. A.J. Christopher traces the quest for a census of the British Empire from 1840 to 1940. He argues that mapping and the census were the two most important enquiries undertaken by the empire. The quest for an imperial census actually starts somewhat earlier, in 1801, when a decennial enumeration of the United Kingdom was instituted and when a question emerged about how to count colonies in this enumeration. How to act at a distance to govern unfamiliar events, places, and people for the empire was eventually resolved with the census as a technology of knowledge production. As Christopher says "as such it represents a significant attempt by the state to number and assess the population and its characteristics and so obtain a view of the society it seeks to govern" (Christopher 2008, 269).

This idea of an imperial census was put forward in the 1840s and remained an objective for a hundred years (Christopher 2008, 271). It followed the introduction in 1800 of "An act for taking an Account of the Population of Great Britain, and the Increase or Diminution thereof", which directed the taking of a census of England and Wales in 1801. It involved the first comprehensive, systematic and centralised collection of information on households and individuals. Previously, numerous local and central institutions and officials linked to the central state, such as ecclesiastical courts, justices of the peace and overseers of the poor, regularly collected information about people in their jurisdictions (Higgs 2004). The census, along with civil registration, replaced that which had been largely carried out by local administrations and dispersed across thousands of archives (parish chests, diocesan registers, estate papers). However, until 1841 censuses did not list individuals but instead provided simple head counts (numbers of men, women, families, and houses) and information about household characteristics such as occupations and ages. The schedules were also completed by officers of the established church or of the poor law system who calculated totals from parish registers. The first nominal census was not conducted until 1841 by the newly established General Register Office (GRO), which initiated "the practice of instructing enumerators to hand out schedules to household heads for them to supply details of the members of their households on

Census night" (Higgs 2004, 72). However, while individuals were counted, unlike parish records the object was the population (Ruppert 2012) where individuals "are no longer pertinent as the objective, but simply as the instrument, relay, or condition for obtaining something at the level of population" (Foucault 2007, 42).

It is in relation to the centralised collection of data on individuals within the British state that the idea of an imperial census was enacted in 1821 with a Colonial Office requirement that each colonial government produce an annual statistical and informational Blue Book (Christopher 2008, 271). The same request specified that population be divided into white, free coloured, and slave categories. After the abolition of slavery in 1833 the categories were simplified into white and coloured. Although by the 1840s virtually all colonial governments had taken annual statistics, each developed divergent practices and produced incompatible annual Blue Books. To solve this fragmentation of information the Colonial Office in London developed the concept of a unified census of the British Empire. With the establishment of the General Register Office and the appointment of William Farr as Superintendent of Statistics, who served from 1839 to 1879, the idea of an imperial census was articulated (Christopher 2008, 272). The General Register Office with the Superintendent of Statistics served as the "centre of calculation" for the analysis, interpretation, collation, compilation, and presentation of the imperial census. This is also when the quest began for dividing the colonies into coherent and comparable enumeration districts by following administrative divisions and boundaries, enumerating both settled and mobile populations, standardising enumeration periods, counting indigenous populations differently, establishing a person's name as a unique identifier, and enumerating not only age and race but also occupation according to categories originally developed by William Farr (Christopher 2008, 273–274).

The quest produced various censuses throughout the rest of the nineteenth century with myriad fits and starts and problems as well as resistances and noncompliances based on different grounds. Colonial Governors were required to explain not only divergences but also resistances. Regrettably, there is no systematic study of these divergences and resistances in the British colonies but an event in 1861 attests to such resistance and perhaps provides evidence of imperial responses. Ceylon avoided conducting the 1861 census and when asked an explanation the Governor simply gave that it was seen as a precursor to taxation, a response that was common in both metropole and colony (Ruppert 2014). Ten years later when Ceylon did conduct a census, its Governor was apparently compelled to reassure the power elite of the colony by stating that

> the Census has no connection whatever with taxation but is taken solely for the purpose of ascertaining the number, ages and occupations of the inhabitants of the island. The information is required in order to ascertain whether or not the population of Ceylon are prosperous and increasing in number, and to enable the Government to devise measures for promoting the improvement of the country and the welfare of the people.
>
> *(quoted in Christopher 2008, 276)*

As Christopher notes

> the quest for a systematic synchronised population census of the British Empire lasted for a hundred years. It represented an attempt by the Colonial Office to obtain a view of the Empire as a whole as an aid to its efficient administration, although the precise use of the census was never explicitly stated.
>
> *(Christopher 2008, 284)*

He concludes that "nevertheless, only one official integrated Report on the Census of the British Empire was ever published" (2008, 284). Christopher thinks two significant factors contributed to this "failure". First, although successive Registrars General of England and Wales periodically monitored the development of the project, the colony was always secondary to metropole. As such, the colony received only limited resources especially for processing the data collected (2008, 284). Second, which is important from our point of view, was that British Empire found it difficult to co-ordinate diverse, multiple, and relatively autonomous colonial governments (Christopher 2008, 284). This is worth investigating further. To what extent colonial governments resisted an imperial census and to what extent this constituted on the part of colonial authorities gaining power-knowledge over their "own" populations are questions that arise from the quest.

Nonetheless, the quest for an imperial census that lasted a century attests to how counting, categorising, and ordering worked as technologies of colonial government and how these technologies were inherited, reshaped, and reused by postcolonial governments. We suggest that the transition from colonial to postcolonial governments involved considerable continuity of technologies of government, especially the use of data for counting, categorising, and ordering. The quest for a unified census of the British Empire between 1840 and 1940 produced vast amounts of data but eventually became a failed project in the sense that James Scott defined the emergence of the state. Even if it eventually failed to accomplish its stated objectives the quest, with its trials, tribulations, ridges and troughs, produced an emerging logic of imperial government. Additionally, the quest led to myriad other intentional or collateral effects such as the invention of new governmental practices, and bureaucratic and technological infrastructures. Furthermore, although Christopher considers the census as "stocktaking" – as governments typically do - it is well to remember James Scott's conclusion (1999, 3): quests of mapping territory or enumerating population were not merely description exercises but are technologies of government that "give [their] categories the force of law". Moreover, it is well to remember Hacking's distinction between overt and "subversive" effects of the census. He says that the overt amassing of gigantic amounts of data rarely results in its intended effects. For Hacking "[t]he fetishistic collection of overt statistical data about populations has as its motto 'information and control,' but it would more truly be 'disinformation and mismanagement'" (Hacking 2015, 281). Yet, Hacking says,

> there is a quite unintended effect of enumerating, and I call this subversive. Enumeration demands kinds of thing; or people to count. Counting is hungry for categories. Many of the categories we now use to describe people are by-products of the needs of enumeration.
>
> *(Hacking 2015, 280)*

He concludes that "biopolitics as the transition from the counting of hearths to the counting of bodies" (which we noted earlier) follows from this. Thus, "the subversive effect of this transition was to create new categories into which people had to fall, and so to create and to render rigid new conceptualizations of the human being" (Hacking 2015, 281).

We see a century of quest to produce a census of an empire therefore not so much as a quest to describe it but as a quest to govern people in a different way. What drives the will to knowledge – the insatiable accumulation of data, classifications (race, ethnicity, language, religion), categorisations (caste, tribe, kin), interpretations (normal, abnormal, deserving, underserving, dangerous, useful), inferences, relations, processes, identifications – is the will to power: a force to maintain, nurture, sustain, and encourage capacities that are useful for the purposes for which there is an interest in the object. The census and its practices – collecting, collating, and presenting data and drawing conclusions about the population as an object – were not only for creating a coherent and consistent method – which always failed – but the will to know that satisfied the will to power. The more difficult it was to establish a coherent or consistent system the more driven was the will to know. Yet this will to know was bound to exceed the technologies that afforded it: the census was increasingly surrounded by and indeed gave rise to various knowledge practices: ethnographies, comparisons, theories, interpretations, disagreements, and debates that veritably began creating an image of the colonial state and society. Census, or the attempts to create an imperial census, generated various forms of data and knowledge as well as agents and authorities who subsequently as experts developed autonomous interests and practices in data that created a condition of possibility for the birth of the postcolony.

The development of the modern census throughout the nineteenth century is often told through methodological nationalism: a survey of nation-by-nation developments where each nation develops its legislative authority and administrative machinery and conceived of populations as contained within national boundaries (Chernilo 2011, Dumitru 2014, Scheel et al. 2016, Wimmer and Glick Schiller 2002). Yet the articulation of census standards is also an imperial development of a different order. Almost at the same time when European imperial states began to develop census practices an international field of data politics emerged where newly developed methods and techniques and their experts began to establish and compete in the development of protocols for international co-operation and standardisation. The First International Statistical Congress was held in Brussels in 1853, which adopted formal international recommendations for conducting a census urging comparability amongst various national censuses.

In 1872, the International Statistical Institute met in St. Petersburg and adopted not only standards for conducting censuses but also methods and data (Goyer and Domschke 1983, 8). When the same institute met in 1897 the idea of a census of the whole world was articulated. Shortly after its formation following the second world war, in 1950, the United Nations (UN) defined one of its urgent tasks as the development of census methods and data standardisation. Practically, this involved tweaking standards developed since the 1850s. It was not until 1970 that the UN developed its World Population Census Program (Goyer and Domschke 1983, 9). Since then the United Nations has progressively developed census guidelines to achieve greater standardisation across all states and regional organisations such as the European Commission have enacted statutes and regulations that comply with and extend these to achieve harmonised European data statistics. Despite these efforts, the interpretation and implementation of standards has been uneven and variable due to a combination of political, technical and historical differences. As we noted in relation to the quest for an imperial census, effects of governmental logics are not reducible to their stated intentions and objectives but give rise to myriad collateral strategies, responses and effects.

Notwithstanding this variability the will to know through censuses has performative effects including the emergence of its agents within a transnational field of statistics (Scheel et al. 2016). The emergence of guilds of experts and star figures of statistics such as William Farr (1807–1883) signals only a glimpse of a veritable structure of expertise about both metropole and colonial populations and the interests of experts that were related yet by no means reducible to the interests of the old European empires and perhaps constituted a different empire – data's empire – that secured and maintained its hold on colonial populations long after they had become postcolonial. As Hacking mused thirty-five years ago "it will be salutary if some of us go on noticing mutations within the more gradual expansion of the biopolitical empire" (2015, 281).

This still remains a hypothesis, as it were, since there are not many studies on the attempt to create a British imperial census let alone relate it to imperial government. But new studies such as those by Karl Ittmann, Dennis Cordell, Gregory Maddox (Ittmann, Cordell, and Maddox 2010) and their colleagues on attempts in Africa and by Jen Emigh, Dylan Riley, and Ahmed Patricia (Emigh, Riley, and Ahmed 2016) in Asia already indicate that Foucault's hypotheses about government and population are bearing fruit – albeit with modifications and enhancements – for studying the imperial and colonial government of populations. We expect that these studies will grow and will form part of what we call "postcolonial data politics" and enable us to better understand the relation between the production of colonial populations through the accumulation of data and imperial government. Meanwhile, we also think there is an urgent need for studies that examine contemporary data politics and its postcolonial implications. We now want to illustrate what we see as emerging areas of investigation in postcolonial data politics, which Hacking calls "biopolitical empire" and which we recast as "data's empire".

Governing postcolonial peoples

It is within the historic amassing of data that we approach the current "deluge" of data (Hey and Trefethen 2003) – which now has come to be named "big data" – to consider what bearing the constitutive powers of knowledge (of population, of territory, and of memory) has for this new era of data politics. The range and volume of data about populations and other objects being generated through the internet and myriad digital devices by organisations, agencies, corporations and governments is unprecedented. But beyond volume, data is also being remade in standardised forms that traverse national borders and with qualities that are increasingly granular, immediate, varied, and detailed. From Facebook claims that it can make a map of everyone in the world (Meyer 2016) and that it "reaches more people than the U.S. Census data says exist" (Swant 2017) to Google Street View being proposed as an alternative for generating census statistics on socioeconomic characteristics (Gebru et al. 2017), there is no shortage of claims to knowledge. Arguably, no kingdom, state, empire, government, transnational or global organisation or corporation has ever held such command over the production, storage and analysis of data. The implications of these developments associated with big data have been mostly debated in relation to privacy, anonymity, security, speech, and other concerns that are cast in terms appropriate to Euro-American metropoles. As we have argued, just as metropole-colony relations were configured through the invention of biopolitics, we expect that new developments in big data are reconfiguring not only metropole-postcolony relations but also biopolitics. To put it differently, not only do the issues of concern such as privacy and security play out very differently in the metropole than in the postcolony but also agents of power-knowledge and their targets of government are also being constituted differently. Moreover, data politics are also playing out differently in the metropole and postcolony precisely because of different trajectories through which certain rights are protected, developed, or violated.

All this is most evident in big data projects in Africa. African postcolonial states are said to face a longstanding "knowledge problem" because of flawed development data on metrics such as Gross Domestic Product (GDP) and economic development (Jerven and Johnston 2015). Some African national statistical departments have more resources and stronger capacity and experience than others, such that development problems include the unevenness and often absence of "good" statistical data (Jerven 2013). It is in this context that big data is imagined as an opportunity to know Africa and other postcolonial states in unprecedented ways, which are critical for decisions about development in areas such as healthcare, security, economic productivity, and disaster and resource management, and so on (Hilbert 2016).

That is the promise reflected in the United Nations Global Pulse initiative, which is principally focused on adopting big data for monitoring and reaching sustainable development goals and managing humanitarian action. The initiative was established based on the recognition that big data offers the opportunity to gain a better understanding of changes in human well-being, and to get real-time feedback

on how well policy responses are working (United Nations 2018). It is "intended as a Call to Action to inspire development agencies and particularly evaluators to collaborate with data scientists and analysts in the exploration and application of new data sources, methods, and technologies" for "programme monitoring, evaluation and learning" (Bamberger 2016, 22). Global Pulse is only one of several initiatives involving projects that engage with either publicly available or big data donated by "large multinational corporations such as Orange and Twitter for purposes of monitoring and evaluating social or economic dynamics in LMICs (Low and Medium Income Countries)" (Taylor and Broeders 2015, 231). The initiative engages with and is part of a broad network of international researchers and organisations working with big data on development projects and who publish, discuss and circulate results, and share knowledge about interventions.

More generally, Global Pulse is part of a "field of study" referred to as "Information Communication Technologies for [international] Development" – ICT4D – described as "an interdisciplinary practice that combines tech with international development, human rights, and public health" to collect, store, process, analyze and share data for development (Anonymous 2016). The kinds of data vary and can include everything from health care data to mobile phone metadata, sensor and biometric data and survey data.

Projects involving big data analytics are undertaken as part of Global Pulse have included numerous experiments with mobile phone data such as the mapping of poverty in China using call data records and mapping population displacement in Nepal following the April 2015 earthquake (Bamberger 2016, 41). Others include using social media to explore HIV-related stigma in Rio, guiding emergency services in the aftermath of the Haiti earthquake, and detecting and managing forest and peat fires in Indonesia. One 2013 project involved the analysis of anonymised mobile phone data to visualise population movements in Senegal (Global Pulse 2015). Through the use of visualisations, a series of mobility profiles were produced for different regions to identify how changes in patterns of mobility could indicate changes in livelihoods or coping strategies, or exposure to new shocks. Monitoring such changes for vulnerable groups in "real time" was identified as potentially offering a "powerful humanitarian early warning mechanism for informed decision-making and rapid response" (Global Pulse 2015, 1). These are just a few amongst a large number of big data projects.

While much has been made about the potential of these various forms of big data to finally "know" postcolonial populations, critics rightly point to the "politics, power dynamics and ongoing patterns of privilege and marginalisation on a global scale" of ICT4D initiatives and especially the lack of ethical processes such as informed consent and opt-out procedures that "continue the legacy of colonialism within aid work" (Anonymous 2016). They argue that many projects in the postcolony lack data protection for personally identifiable data, which would be unacceptable in the metropole where the governments or corporations that own and control data are located. These practices mean

> there is a danger of setting up a form of imperialism based on personal data. Just as the royal powers of old reached far into the lives of distant colonised people, technology companies gain immense control with every terabyte of personal data they store and analyze.
>
> *(Simmons 2015)*

For Simmons, technology companies that are predominantly owned and located in the United States are colonial in their actions as they perform like sovereign nations and increasingly operate across borders.

The reach of these practices, however, extends to both metropole and postcolonial populations and thus they need to be understood, as we argued, in relation to each other. That is, technology companies operate transnationally to harvest the data people generate in their day-to-day lives through a form of "accumulation by dispossession that colonizes and commodifies everyday life in ways previously impossible" (Thatcher, O'Sullivan, and Mahmoudi 2016, 990). Thatcher et al. interpret this through the metaphor of "data colonialism" to highlight "the power asymmetries inherent in contemporary forms of data commodification." However, just as censuses and other forms of state produced data were uneven in their colonisation, big data are leading to a "new kind of digital divide" in "data-based knowledge" due to the unevenness of technological diffusion as a result of "lack of infrastructure, human capital, economic resource availability and institutional frameworks" (Hilbert 2016, 135). Beyond these "contextual" variables, just as in colonisation, the varying assemblages of experts, methods, technologies, data, organisations, guilds, associations, practices, authorities and other interests are constitutive of uneven effects beyond the control of the technology companies.

For these reasons, while these critiques are apt the situation is more complex. Importantly for our argument, they do not capture what we suggest is the continuity and discontinuity from the regulatory logic of biopolitics that characterised data politics of the last two centuries and was concerned with population as a species-body. We thus refer to "data's empire" to signal an emergent regime of government that involves new as well as existing mechanisms of domination between the metropole and postcolony that is producing a species-body with different characteristics and with heterogeneous effects. These characteristics are yet to be fully investigated. What we offer in this chapter is thus a preliminary outline of emergent mechanisms that are distinct and overlapping and have continuities and breaks from past ones. We then discuss their potential implications for the biopolitics of our present.

The species-body is re-assembled. Numerous actors and arrangements (technologies, practices, data, methods, agencies, authorities, professions, and so on) make up data's empire. Beyond the "public–private partnerships" and "growing agency of corporations as development actors" (Taylor and Broeders 2015, 229), various combinations of actors that traverse both national borders and public-private sectors are engaged in the production of the species-body. They involve states, agencies, organisations, corporations, and institutions that come together

to process data for different purposes of government. ICT4D projects, for example, are funded by governments and nongovernmental organisations (e.g., UN), corporations, private foundations, and private individuals. Many involve collaborations between government and non-government organisations and various professions such as statisticians and data scientists. Examples include a project that analysed financial transaction data of the BBVA bank to measure the economic resilience of populations to natural disasters in Mexico and another used satellite imagery data produced by various governments to track poverty trends in Uganda (Bamberger 2016). Significantly, the agents of these assemblages occupy relative positions within the transnational field that includes professionals such as data scientists, statisticians, programmers, software developers, methodologists and so on who traverse transnational and national borders (Scheel et al. 2016). Rather than state authorities being replaced or superseded by corporations or private sector actors, the fields of power and knowledge and the agents and practices that make them up are being diversified and recomposed. Furthermore, numerous projects such as those of Global Pulse are not separate from but part of this transnational field that engages with big data in the metropole. Just as a quest for an imperial census happened simultaneously to the development of national modern censuses, so too are quests to know populations through big data happening transnationally.

This is one reason we suggest that the species-body is re-assembled. Its production includes various combinations of technologies such as the hardware of computers but also sensors, satellites, antennas and mobile devices and beyond software to include analytics such as algorithms, AI, machine learning, and cloud computing. These different combinations of sources and technologies traverse borders and fields as do the data they produce. Mobile phone data, for example, is produced in myriad contexts and taken up and combined with other data to enact phenomena such as migration flows, disaster responses, and economic well-being (Tazzioli 2018). More generally, millions of data points can be assembled on several hundred variables related to a topic of interest such as an individual's transactions, weather patterns, and social media postings. Many data sources can be brought together into a "high-dimensional space" that envelopes and flattens differences between data (Mackenzie 2015, 434). In these ways, rather than replacing existing data regimes, myriad data are combined from a "continuum" that includes "big data" (digital transactions, social media, sensors, etc.), "large data" (census, survey, administrative, etc.) and "small data" (qualitative interviews, focus groups, etc.) (Bamberger 2016, 39).

So, while the imperial quest for a census sought to colonise through standardised and pre-defined methods and data organised and managed principally by imperial states, data's empire consists of a proliferation of data that is produced in ways that are dispersed and distributed (Ruppert, Law, and Savage 2013). Data can move and circulate between different sites of production and be "repurposed", that is, used for generating knowledge to serve governing objectives sometimes far removed from that for which the data were originally produced. Call data records of mobile phone operators for instance are repurposed to know changes in the

livelihoods of a population and can be done so across spaces not confined to territorial borders and in relatively standardised, comparable and interoperable formats. That is, the data of the transnational field is standardised in ways not previously attained or even attainable by previous data regimes.

The species-body is multiple. This multiplication does not merely constitute "new" representations of "old" populations. The multiplication of assemblages also multiplies the object, the species-body. To understand the species-body as multiple is to first understand data as a performative entity: it does not merely describe but produces the objects it represents. It is for this reason data can be understood as an actant within assemblages for what they might perform changes depending on the relations they enter and through which they have agentic qualities. What data perform is brought into being by assemblages of experts, methods, technologies, data, organisations, guilds, associations, practices, authorities, and other interests but it is never under the strict control of any of them. It is through its circulation and repurposing that data are detached from the assemblages that make them up and come to act on and colonise objects as well as subjects in myriad ways.

The species-body is performed. Multiplication engenders subversive effects. If the massive efforts of the British Empire over a century failed to produce an imperial census it did succeed in producing colonial populations and institutionalising ways of representing, measuring, counting, and acting on them. And as Hacking argued, a subversive effect of such enumerations was the making of categories that rendered "rigid new conceptualisations of the human being." The subversive effects of data's empire today supplement this logic of categorisation as a governing strategy to colonise individual and population bodies but differently. Conventional population statistics typically involve sociodemographic categories and then collecting data through usually self-elicited accounts that use various methods to fit people into them. While this logic persists, the repurposing of big data involves analytics that identify categories and classifications of populations rather than imposing them in advance. Categories on mobility or economic health are generated as a consequence of analytics such as machine learning that do not identify associations between a limited set of existing variables, but explore multidimensional patterns amongst "hundreds and in some cases tens of thousands of variables and sample sizes of millions or billions data" (Mackenzie 2015, 434). Differences are not understood as "variables" as in classical statistics, but derived from combinations of attributes or "features" from myriad "forms of data (text, images, video, transactions, sensors), not just the variables measured using classical statistical tabulations of surveys, polls or random sampling" (Mackenzie 2015, 433). For these reasons, Mackenzie argues these analytics involve a different mode of knowing differences through classification, which involves "the generalization of prediction". That is, while predictive modelling is not new and indeed part of all regimes of power/knowledge, the innovation is its expansion to incorporate not only large volumes of data but a wide range of features or attributes (e.g., transactional data, social media posts, weather readings) within a generalised space to find "useful approximations" (Mackenzie 2015, 435). In development projects, this is

expressed as the creation of "standardized data categories into which many different types/sets of information can be fitted so that data are comparable over time and space" (Bamberger 2016, 38).

While features or variables can be diverse, a key difference from the conventional statistical production of population is the registering of multiple forms of conduct or what people do such as their movements and actions (transactions, choices, statements, interactions) where the focus of inquiry is not on the individual factors that affect conduct, but on aggregate patterns and connections: contagion, influence, association, etc. (Ruppert, Law, and Savage 2013). It is the continuous tracing of the conduct of the species-body that analytics of mobility, sentiment or transactions are based. The species-body is thus a performing body that is not stable but always becoming. For example, in the Senegal project, an algorithm grouped individuals within different livelihood zones according to their movement trajectories throughout a year to reveal distinct mobility patterns and groupings (Bamberger 2016, 1).

The species-body is visualised. The species-body is not only multiple and enacted. It is also visualised. As Edward Tufte (1983) famously insisted, visualisation can be a technique "for reasoning about statistical information" (9) that 'reveals data' and can be more 'precise than conventional statistical computations' (13). Now with millions of data points, visualisation has become a key technique of making data visible and forms "part of the toolkit that data miners and data scientists employ to navigate, transform or otherwise explore data" (Mackenzie 2015, 437).

Visualisations can identify patterns previously unseen and include interactive elements and dashboards that enable seeing the effects of combining different data on features of a population (Bamberger 2016). In the Senegal project, data on monthly rainfall for each livelihood zone could be incorporated at "different geographical and temporal resolutions using remote sensing data from NASA" to visualise the impact on population movements (Global Pulse 2015, 2). Beyond representation, visualisation is thus also an analytic that makes it possible to detect and observe the species-body. Unlike the maps documented by Anderson, visualisation reveals populations as patterns, trends and tendencies immanent in the species-body as a changing being.

The species-body is alive. Combined together, the multiplicity, performativity, agency and subversive effects of data render it uncontrollable by a central authority yet manageable because of the possibilities of detecting the species-body and then calibrating the conduct of the individual body. It is this last aspect of data that best signifies what we mean by "data's empire". Rather than the periodic "stocktaking" of conventional statistics, populations are living bodies that have pulses, flows, and patterns. In turn, data serves a dual function: for identifying attributes or features (e.g., sentiments about the economy) but then monitoring and evaluating those features over time (e.g., daily changes in sentiments about the economy) and then intervening through specific governing projects. The former engages data in much the same way as classical data regimes: populations as entities to be measured. However, the latter is what big data makes possible: the potential to monitor and

evaluate the performance of the species-body on a more continuous basis. It is a logic captured in four stages of data analytics: descriptive and exploratory analysis (what is happening, often in real-time); doing predictive analytics ("what is likely to happen"); detection ("tracking who is likely to succeed and who will fail"); and evaluation and data diagnostics ("how to improve programme performance") (Bamberger 2016, 60–61). What these four stages capture is how data performs the relation between sovereign, disciplinary and regulatory logics. In the Senegal project, rather than evaluating using historic survey data, real-time information of trends and changes in mobility provide "early warning of emerging vulnerabilities" thereby enabling rapid response (Bamberger 2016, 1). Response then is at the project level and involves humanitarian interventions to address shocks to livelihoods, for example. In other cases, it can involve targeting conducts (mobility, violence, education) to discipline and regulate individual bodies. Furthermore, the relation between stages is understood as cyclical and involves complex relations and "feedback loops" between monitoring, evaluation, and interventions. In relation to prediction, Mackenzie (2015) notes that big data analytic models are based on the assumption that relatively stable classifications exist. Yet, the performativity of prediction means that these are mutable and thus models must frequently be changed to maintain their predictive power (Mackenzie 2015, 442). For these reasons, the species-body is not measured but calibrated where governing programmes and projects need to be "smart, agile and adaptive" (Bamberger 2016).

Decolonising data's empire

When placed within both longer and shorter series of developments in biopolitics and its allied technologies (statistics, demography, census), more recent developments in predictive analytics, algorithms, machine learning, and the like begin to appear in a different light. The latter are not recent or simply technical developments but belong to a series of long- and short-duration transformations since the end of the eighteenth century that inaugurate, supplant, and supplement ways and logics of governing peoples. It is important to place these ways and logics into historical series to understand what is enduring and what is changing. If indeed modern logics of governing the Euro-American metropoles were implicated in governing colonial populations in ways that we have yet to understand, how is the species-body performed today by a combination of technologies of big data, large data, and small data? If the species-body is now reassembled, multiple, performed, visualised, and alive in ways that were inconceivable a generation ago how does this inaugurate new logics of governing peoples? How does it reconfigure metropole-colony and metropole-postcolony relations? And in turn, with what governing effects?

Like the "failed" British imperial census, the will to know the species-body of the world by organisations such as the United Nations and its regional commissions experience resistance through the uneven interpretation and implementation of standards, and varying practices and data production regimes in both the metropole and the postcolony. In part this is a consequence of practices being caught within

national approaches and contexts and their variable capacities and investments. What does this mean for the regulatory logic of biopolitics and with what subversive effects for how people are categorised and how populations are constituted as objects of governing? We articulated this question through a discussion of some contemporary practices of international organisations and how they are mobilising big data to address limitations in past efforts to produce data about and to know postcolonial populations. We argue that the continuity between these efforts and quests to do an imperial census of the British Empire becomes more prominent and apparent in postcolonial states that have been drawn into the orbit of "development" discourses with the increasing involvement of technology companies that both spur and produce data needs for government. Unlike conventional empires, however, which are still in operation through various practices, data's empire functions through assemblages of actors, arrangements, technologies, and logics that are transversal: neither their practices nor objects of government are confined or limited a given territory and its security. That the species-body is multiple, performed, visualised, and alive means it has acquired qualities that, unlike conventional vital statistics, measure vital signs of populations, as the increasing use of the terms "pulse" and "sense" indicate.

We are not yet ready to name and add these mechanisms as a fourth regime to Foucault's trilogy – sovereignty, discipline, and regulation. We suggest these mechanisms indicate an emergent logic of how population knowledge about the species-body is being enacted. While these mechanisms have been variously identified by other researchers, we have sought to bring them together in relation to identifying them as part of an emerging logic. We don't think this logic supplants or displaces sovereignty, discipline, or regulation but it is a key logic that is emerging within contemporary data regimes. For us data's empire is unlike early modern and modern Euromerican empires yet it inherits logics of government from them and institutes these new mechanisms of power and principles of knowledge.

References

Anderson, Benedict R. 2006. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. 2nd ed. London: Verso. Original edition, 1983.

Anonymous. 2016. "Data Colonialism: Critiquing Consent and Control in 'Tech for Social Change.'" *Model View Culture*, 43. Accessed 14 November 2017: https://goo.gl/wQSkoV.

Appadurai, Arjun. 1993. "Number in the Colonial Imagination." In *Orientalism and the Postcolonial Predicament: Perspective on South Asia*, edited by Carol Breckenridge and Peter Van der Veer, 314–339. Philadelphia, PA: University of Philadelphia Press.

Bamberger, Michael. 2016. *Integrating Big Data into the Monitoring and Evaluation of Development*. New York: United Nations Global Pulse.

Bourdieu, Pierre, and Loïc J. D. Wacquant. 1999. "On the Cunning of Imperialist Reason." *Theory, Culture & Society* 16 (1): 41–58.

Burbank, Jane, and Frederick Cooper. 2010. *Empires in World History: Power and the Politics of Difference*. Princeton, NJ: Princeton University Press.

Chernilo, Daniel. 2011. "The Critique of Methodological Nationalism: Theory and History." *Thesis Eleven* 106 (1): 98–117.

Christopher, A. J. 2008. "The Quest for a Census of the British Empire C.1840–1940." *Journal of Historical Geography* 34 (2): 268–285.

Christopher, A. J. 2009. "Delineating the Nation: South African Censuses 1865–2007." *Political Geography* 28 (2): 101–109.

Cohn, Bernard S. 1996. *Colonialism and Its Forms of Knowledge: The British in India*. Princeton, NJ: Princeton University Press.

Cordell, Dennis D., Karl Ittmann, and Gregory Maddox. 2010. "Counting Subjects: Demography and Empire." In *The Demographics of Empire: The Colonial Order and the Creation of Knowledge*, edited by Karl Ittmann, Dennis D. Cordell, and Gregory Maddox, 1–21. Athens, OH: Ohio University Press.

Curtis, B. 2004. "The Re-Cutting of Lower Canada in the 1830s: An Essay on Colonial 'Governmentality.'" *Revue D Histoire De L Amerique Francaise* 58 (1): 27–66.

Dumitru, Speranta. 2014. "Qu'est-Ce Que Le Nationalisme Méthodologique? Essai de Typologie." *Raisons Politiques* 2014/2 (54): 9–22.

Emigh, Rebecca Jean, Dylan J. Riley, and Patricia Ahmed. 2016. *Changes in Censuses from Imperialist to Welfare States: How Societies and States Count*. Vol. 2. Basingstoke: Palgrave Macmillan.

Foucault, Michel. 1978. *The History of Sexuality: An Introduction*. New York: Pantheon Books.

Foucault, Michel. 2003. *Society Must Be Defended: Lectures at the Collège de France, 1975–76*. Translated by D. Macey. Edited by M. Bertani. New York: Picador.

Foucault, Michel. 2007. *Security, Territory, Population: Lectures at the Collège De France 1977–1978*. Basingstoke: Palgrave Macmillan.

Gebru, Timnit, Jonathan Krause, Yilun Wang, Duyun Chen, Jia Deng, Erez Lieberman Aiden, and Li Fei-Fei. 2017. "Using Deep Learning and Google Street View to Estimate the Demographic Makeup of Neighborhoods across the United States." *Proceedings of the National Academy of Sciences*. 114 (50): 13108–13113.

Gilroy, Paul. 2004. *After Empire: Melancholia or Convivial Culture?* London: Routledge.

Global Pulse. 2015. "Analysing Seasonal Mobility Patterns Using Mobile Phone Data." UN Global Pulse. http://www.unglobalpulse.org/sites/default/files/UNGP_ProjectSeries_Mobility_Senegal_2015_0.pdf.

Goyer, Doreen S., and Eliane Domschke. 1983. "Introduction." In *The Handbook of National Population Censuses. Latin America and the Caribbean, North America, and Oceania*, edited by Doreen S. Goyer and Eliane Domschke, 3–31. Westport, CN: Greenwood Press.

Hacking, Ian. 2015. "Biopower and the Avalanche of Printed Numbers." In *Biopower: Foucault and Beyond*, edited by Vernon W. Cisney and Nicolae Morar, 65–80. Chicago, IL: University of Chicago Press. Original edition, 1983.

Hevia, James Louis. 2012. *The Imperial Security State: British Colonial Knowledge and Empire-Building in Asia*. Cambridge: Cambridge University Press.

Hey, Tony, and Anne Trefethen. 2003. *The Data Deluge: An E-Science Perspective*. New York: Wiley.

Higgs, Edward. 2004. *The Information State in England: The Central Collection of Information on Citizens since 1500*. Basingstoke: Palgrave MacMillan.

Hilbert, Martin. 2016. "Big Data for Development: A Review of Promises and Challenges." *Development Policy Review* 34 (1): 135–174.

Ittmann, Karl, Dennis D. Cordell, and Gregory Maddox. 2010. *The Demographics of Empire: The Colonial Order and the Creation of Knowledge*. Athens, OH: Ohio University Press.

Jerven, Morten. 2013. *Poor Numbers: How We Are Misled by African Development Statistics and What to Do About It*. Ithaca, NY: Cornell University Press.

Jerven, Morten, and Deborah Johnston. 2015. "Statistical Tragedy in Africa? Evaluating the Data Base for African Economic Development." *The Journal of Development Studies* 51 (2): 111–115.

Kalpagam, U. 2000a. "Colonial Governmentality and the 'Economy'." *Economy and Society* 29 (3): 418–438.

Kalpagam, U. 2000b. "The Colonial State and Statistical Knowledge." *History of the Human Sciences* 13 (2): 37–55.

Kalpagam, U. 2001. "Colonial Governmentality and the Public Sphere in India." *Journal of Historical Sociology* 14 (4): 418–440.

Kumar, Krishan. 2010. "Nation-States as Empires, Empires as Nation-States: Two Principles, One Practice?" *Theory and Society* 39 (2): 119–143.

Kumar, Krishan. 2017. *Visions of Empire: How Five Imperial Regimes Shaped the World*. Princeton, NJ: Princeton University Press.

Ludden, David. 1993. "Orientalist Empiricism: Transformations of Colonial Knowledge." In *Orientalism and the Postcolonial Predicament: Perspective on South Asia*, edited by Carol Breckenridge and Peter Van der Veer, 250–278. Philadelphia, PA: University of Philadelphia Press.

Mackenzie, Adrian. 2015. "The Production of Prediction: What Does Machine Learning Want?" *European Journal of Cultural Studies* 18 (4–5): 429–445.

Mahler, Anne Garland. 2018. *From the Tricontinental to the Global South: Race, Radicalism, and Transnational Solidarity*. Durham, NC: Duke University Press.

Mbembe, Achille. 2001. *On the Postcolony, Studies on the History of Society and Culture*. Berkeley, CA: University of California Press.

Meyer, Robinson. 2016. "Facebook Is Making a Map of Everyone in the World." Accessed 24 July 2018: https://bit.ly/2LilEQt.

Muldoon, James. 1999. *Empire and Order: The Concept of Empire, 800–1800*. Basingstoke: Macmillan.

Pagden, Anthony. 2001. *Peoples and Empires*. London: Weidenfeld & Nicolson.

Ruppert, Evelyn. 2012. "Seeing Population: Census and Surveillance by Numbers." In *Routledge International Handbook of Surveillance Studies*, edited by K. Ball, K. Haggerty, and D. Lyon, 209–216. London: Routledge.

Ruppert, Evelyn. 2014. "Infrastructures of Census Taking." In *The Dawn of Canada's Century: Hidden Histories*, edited by Gordon Darroch, 51–77. Montreal: McGill-Queen's University Press.

Ruppert, Evelyn, John Law, and Mike Savage. 2013. "Reassembling Social Science Methods: The Challenge of Digital Devices." *Theory, Culture & Society*, Special Issue on The Social Life of Methods 30 (4): 22–46.

Said, Edward W. 1994. *Culture and Imperialism*. New York: Vintage.

Said, Edward W. 2003. *Orientalism*. 2nd ed. New York: Vintage. Original edition, 1978.

Scheel, Stephan, Baki Cakici, Francisca Grommé, Evelyn Ruppert, VIlle Takala, and Funda Ustek-Spilda. 2016. *Transcending Methodological Nationalism through Transversal Methods? On the Stakes and Challenges of Collaboration*. ARITHMUS Working Paper No. 1: Goldsmiths University of London. Available at: http://bit.ly/2lqR1aM.

Scott, David. 1995. "Colonial Governmentality." *Social Text* 43 (Autumn): 191–220.

Scott, James C. 1999. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press.

Simmons, Anjuan. 2015. "Technology Colonialism." *Model View Culture* (27).

Steinmetz, George, ed. 2013. *Sociology & Empire: The Imperial Entanglements of a Discipline*. Durham, NC: Duke University Press.

Stoler, Ann Laura. 1995. *Race and the Education of Desire: Foucault's "History of Sexuality" and the Colonial Order of Things*. Durham, NC: Duke University Press.

Swant, Marty. 2017. "Facebook Claims It Reaches More People Than the U.S. Census Data Says Exist." *Adweek*. Accessed 24 July 2018: https://bit.ly/2mBMZ1k.

Taylor, Linnet, and Dennis Broeders. 2015. "In the Name of Development: Power, Profit and the Datafication of the Global South." *Geoforum* 64: 229–237.

Tazzioli, Martina. 2018. "Spy, Track and Archive: The Temporality of Visibility in Eurosur and Jora." *Security Dialogue* 49 (4): 272–288.

Thatcher, Jim, David O'Sullivan, and Dillon Mahmoudi. 2016. "Data Colonialism through Accumulation by Dispossession: New Metaphors for Daily Data." *Environment and Planning D: Society and Space* 34 (6): 990–1006.

Tufte, Edward R. 1983. *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

United Nations. 2018. "United Nations Global Pulse: Harnessing Big Data for Development and Humanitarian Action." Accessed 30 July 2018: https://bit.ly/2GQgqpV.

Wilson, K. 2011. "Rethinking the Colonial State: Family, Gender, and Governmentality in Eighteenth-Century British Frontiers." *American Historical Review* 116 (5): 1294–1322.

Wimmer, Andreas, and Nina Glick Schiller. 2002. "Methodological Nationalism and Beyond: Nation-State Building, Migration and the Social Sciences." *Global Networks* 2 (4): 301–334.

Young, Robert. 2016. *Postcolonialism: An Historical Introduction* 2nd ed. London: Routledge. Original edition, 2011.

PART IV
Rights

12

THE RIGHT TO DATA OBLIVION

Giovanni Ziccardi

Introduction: accumulation of digital data in the information society

In a modern (and connected) information society, the "life" (and "death") of digital data (Moreman and Lewis 2014), whether it involves information collected about a person's online activities (persisting even *after* his or her death), comments posted in memory of a missed friend (Parker 2014; Wilmot 2016), tweets of condolences (Weiser 2016) and chat logs of conversations, is increasingly at the center of legal and social issues and public debates. The right to control, erase, or move "digital footprints" collected through free and paid services has definitely become a central political and legal issue (Steinhart, Nagasawa, and Wielenberg 2014).

Deleting "digital footprints", especially information or news that a person would like to remove from the Internet by pursuing a "right to oblivion", is increasingly a matter of data politics, and is now evident in international legal and political debates. These debates, incidentally, are fully connected to the modern idea of "algorithmic society", which features "large, multinational social media platforms that sit between traditional nation states and ordinary individuals, and the use of algorithms and artificial intelligence agents to govern populations" (Balkin 2018, 1151).

The ability to remember, and to communicate a great deal of information, is what enables humans to connect their past and present and project themselves into the future. Arguably, human civilization is made possible by the ability to transmit memories, from one person to another and from one generation to the next (Malone 2013). The Internet has accelerated the possibilities of preserving, communicating and transmitting information and knowledge. Today the Internet allows people to take advantage of a wide variety of digital sources, or to become creators of content. At the same time, it is clear a great capacity for preserving

digital information, combined with the wide possibility of the creation and communication of content offered by the Internet and the ability to search and analyze data through complex algorithms. The provision of such research possibilities to each user as in the case of search engines, usher a paradigm shift and a revolution in the information society (Floridi 2014).

Such data has already been defined, by scholars and in legislative drafts that aim to regulate the phenomenon, as "digital assets" that have a real value (a *commercial* or an *emotional* value, or both). Referring to the financial/commercial interests involved, recent studies commissioned by IT security companies rated a digital asset (for a typical online user) at the significant sum of $35,000 *per capita* (Siciliano 2013). In the opinion of many, such evaluation requires a new specific regulation far different from rules for material goods. One of the first legal issues has concerned, for example, the value of commercial video and audio files downloaded by users and stored in the computer of a deceased person (Wong 2013). In the United States of America, more than thirty States are issuing – or have already enacted – rules aimed at solving the post-life problem of accounts, profiles and data (Herrera 2017). In Europe, these concerns have been the focus of the promulgation of the General Data Protection Regulation (GDPR) (European Regulation 2016/679, which entered into force on 24 May 2016 and implemented in all Member States from 25 May 2018), together with the possibility of processing similar information and related rights in an automated manner (Wachter, Mittelstadt, and Floridi 2017).

Legal and technology experts agree that legislative regulation of the life cycle of digital data needs to be focused on two facets of the problem:

i) the information technology side, which includes data security, data protection and the best technical approaches to preserve the information for the use of future generations (Schneier 2008), and
ii) the legal side, that involves property rights related to digital assets, the privacy of the deceased (seen as the last opportunity to protect a person's confidentiality, or to keep secret all information that he or she has not intented to reveal) and the heirs' rights to know certain information they need to claim an inherited property or right (Moreman and Lewis 2014; Steinhart, Nagasawa, and Wielenberg 2014).

Unfortunately, control of the life and death of digital data and information in today's hyper-connected societies – especially related to very popular social network platforms – has turned out to be a very thorny, and particularly difficult, subject to handle (Carroll and Landry 2010). First of all, an algorithmic society, as Balkin refers to it, permits the collection of vast amounts of data about persons, and facilitates new forms of surveillance, control, discrimination and manipulation, both by governments and by private companies. Balkin calls this the "problem of Big Data" (Balkin 2018, 1153). What is becoming increasingly recognised by scholars, technicians, politicians and users is that people can exercise ever diminishing control of the life and death of the vast amount of data about them that is generated daily, and that it has become impossible to follow the "destiny" of their

data once generated or digitized and put online. Technological precautions can limit potential harms but, at the same time, it is clear that, in most situations, it is not possible to control the life (or death) of data that circulates about a particular object or person (Garber 2016).

Moreover, the Internet and social networking platforms are moving relentlessly into the creation of two, three and more profiles related to a person (Hildebrandt and Koops 2010), by joining and correlating information (and interpreting traces) that are left behind during daily digital activities. This implies that a person has one or more "electronic bodies" existing online, being updated and, often, beyond their control. The notion of "electronic body" or "electronic person" was elaborated, among others, by the Italian jurist Stefano Rodotà (Rodotà 2015). Internet platforms operate and create data even when a person is offline. These activities make the profile of people's "living clones" (or "avatars") on Facebook, Twitter, Instagram, blog, chat and WhatsApp groups more and more accurate and similar to "real" individuals (Lafrance 2016).

People who have profiles on different services that are, however, very similar in content, are very common: they connect, for example, accounts and update activities on Facebook, Twitter and Instagram to allow, with only one action/post, the automated update of all profiles (Carroll and Romano 2010). At the same time, users who customize each single account and who publish only a certain type of photo on Instagram and report different information depending on whether they are on Twitter or Facebook are now very common too. In the latter case, by carefully analyzing the type of text used, and the tone and syntax of the conversations, persons can be distinguished from one another. This means that often a person's digital side is much more dynamic and subject to variations and as such "liquid", following the evocative definition of Bauman (2000). On the other hand, technology works more and more independently of a person's instructions, and generates new information about the character of the digital person through algorithms that learn, assimilate, re-elaborate and publish data with greater accuracy to profile the user.

Digital data are generated, over time, and technologies work to control its longevity. The first consideration that this gives rise to is how important is the need for a death of digital data to bring back control of information to the person. If indeed big data is seen as the new oil, its life and death pose big technological, social and epistemological challenges (Floridi 2012). Balkin is particularly clear about how the development of algorithms and artificial intelligence play a role in these challenges because "the Algorithmic Society depends on huge databases that can cheaply and easily be collected, collated, and analyzed", and the digital age

> makes all of this possible because digital communication involves creating data, copying it, storing it, and moving the copies from one place to another. In the digital age, more and more things that people say and do leave digital traces that can be collected, copied, collated and analyzed.
>
> *(Balkin 2018, 1155)*

It is perhaps fair to say that today, in many cases the problem of digital social networks is no longer to *remember*, but is to *forget*. Often, operations, decisions and choices that used to be left to human judgement, are now deferred to algorithms (Mittelstadt et al. 2016), with several technological and ethical implications (Floridi 2014). What then are the possibilities of legal and political intervention that can give control to people over the right to the life and death of data about themselves? Any such intervention on the fate of digital data and its oblivion after the death of a person would have to involve an adequate understanding of the nature of the digital data, its protection, its relationship with the privacy of the individual and, above all, the specific workings of social networks through which digital data is disseminated, transformed, correlated and resistant to erasure.

The life and death of digital data and its oblivion thus combines political, technological and legal issues to generate very complex situations. Political issues have become especially critical given both the economic and social power that social networking and other platforms have acquired, and the possibility that national and international laws could actually regulate them in the face of such power. Additionally, the capacity to trace and control digital data about ourselves is every declining as a result of the workings of automation and algormithms. Taken together it is very challenging to propose a framework that could enable tracing and controlling a person's digital data. In response to this challenge, I shall first outline some of the complexities of the life and death of digital data. Then I will focus on data portability and data oblivion as two legal issues that frame this as data politics.

The life and death of digital data

The online activities of a person involve, as a first consequence, the generation of digital data about them such as that concerning their physical, economic or social characteristics. This data can be created either by the person or the workings of digital technologies and platforms through, for example, the processing of data, correlations between data or the profiling of a person's activities. According to a Global Web Index survey of 2015,[1] it was estimated that an average Internet user has about six social media accounts or profiles on different Internet services (the survey analyzed the presence of users in fifty social media platforms). If we add these six accounts to other social network accounts or to typical and common profiles that connect an individual to cloud services, e-mail accounts, home banking credentials and online video providers, the number of these profiles can easily reach twenty per person. Day after day a wealth of data is generated by these services and is moved, changed and processed. Slowly over time these activities generate capital, which becomes interconnected to social life and whose accumulation or lack thereof can generate new problems (cultural, economic, political, social). Often, users do not realize that, in everyday life, Internet services can also, at the discretion of providers, put a real "death penalty" to such digital data. All services are provided with well-defined contractual

terms that most users rarely read carefully, and which in several cases specify the right of the service provider to "abandon" the user, i.e. to suspend or, worse, to cancel the service. Usually, in paid services – like, for example, the activation of a certified e-mail box or a cloud service - such a possibility is slightly mitigated by requiring prior notice, or granting a short amount of time to recover all data thus enabling the user to store them elsewhere.

For "free services" with millions, if not billions, of users, providers have wide margins of discretion. It could be sufficient for the provider to believe that a contract rule has been violated in order to suspend the service and delete all the data. However, among other issues, often erasing is not complete as data remains on backup computers or in a provider's archives such that the data will continue to live without the knowledge of the user who, on the contrary, believes them to be erased.

It is important, therefore, to recognise that the digital data created by the investments, attention and time of users live under the threat that could, from time to time, put an end to them and yet may continue to live without their knowledge. The responsibility is placed on the user who has to read user agreements carefully, be prepared, and must make regular backups and retrieve data when a service is terminated. Many platforms provide a plethora of complex procedures by which users can download, refine, tweak or delete data and can lead users to believe that they are doing all these things whilst it is impossible to know how and where the data continues to live.

There are, however, services that die a "natural death", namely because they are suddenly out of the market, overtaken by competitors or bought by other companies in the same sector and then ended (a very common procedure in technological services). Recall, for example, the booms of Second Life, Orkut, Geocities: all services that had channeled millions of users, data, information, activities and "lives" and that, from one moment to another, have ceased to exist or have been significantly reduced in their importance and size.

How then might we conceive of the possibilities of controlling digital data after the physical death of the person? Given that this data can continue to generate tangible economic and cultural value and generate further information through the workings of social networks, algorithms and the correlation of big data, how can this value be connected to the rights of the person?

In response to these complex and dynamic questions, I identify two "legal-informatics" issues: the right to data portability and the right to be forgotten, as fundamental challenges of data politics today.

A first legal-informatics issue: data portability

When a service dies, the problem becomes that of *data portability*, namely to ensure that there is a possibility of continuing the digital lives of users in another IT system. In the new GDPR, there is an Article about the portability of personal data for those services where data are accumulated over several years, such as a mortgage, loan or e-mail account. It establishes the right to have the data in an orderly, structured and

machine-readable format and in a processable form to allow them a later life and not a technological death.

Considerando 68 of the Regulation, with the intention of illustrating the general principles underlying the idea of portability, clarifies this right, which is directly linked to the life of the data in the future, and which is then reaffirmed in Article 20 of the Regulation itself. Thus, in the first illustrative part of the Regulation (that consisting in 173 recitals/"Considerando"), we have the indication of the importance of the principle, which is then made as a rule in the specific Article of the Regulation itself. The idea of a right to data portability is very similar to the one that underlies the portability of a cell phone number, enabling a person to choose to keep the same number and "take it" to another provider that, for example, offers cheaper conditions. The new right to portability intends to promote data subjects' control over their personal data, and facilitate the circulation, copying or transmission of data from one information technology environment to another.

In order to further strengthen control over data, it is seen as appropriate for the person to have the right, if the personal data are processed by automated means, to receive in a structured, common, machine readable format personal data that they have previously provided to a data controller thus enabling them to transmit the data to another controller. Yet, this Regulation cannot deal with the digital death of data that depend heavily on service, market trends and the technologies and formats used (Moreman and Lewis 2014). It is no coincidence that, after more than thirty years of mass network connectivity, there are many data that are related, in terms of their lives or death, to a service and a format or, conversely, have already been orphaned. In some instances, these services are not orphaned due to the physical death of the user but by the logical death of the service that hosted them or because they are now incompatible with current formats, technologies and standards.

A second legal-informatics issue: the right to be forgotten

While data portability is related to the right to control the life of one's data, a second major challenge for jurists and technicians is the question of how to guarantee the right to forget digital data in a digital context that, on the contrary, does everything to keep data alive. One issue is that the right to "oblivion" brings into conflict the right to free information and free speech, privacy, press rights and right to knowledge, with the right of people to have some aspects of their lives forgotten thereby enabling them to start again (Brock 2016). If the right to oblivion is recognized as part of respecting and protecting fundamental rights in an information society that tends to shape and profile each person, then legal measures are necessary to guarantee it at least to a certain level, since a technical forgetfulness is impossible to be completely assured.

From a legal point of view, the right to oblivion was stated in 2014 by the judgment of the Court of Justice of the European Union in the "Google Spain" case (Case C-131/12, decision of 13 May 2014). Briefly, the right to oblivion, or "right to be forgotten", is "a feature of European data privacy law" (Balkin 2018, 1201)

that was extended to online search engines in 2014 by the Court of Justice of the European Union. Balkin recalls how

> the case arose out of a 2010 complaint to the Spanish Data Protection Agency by a Spanish lawyer, Mario Costeja González. Costeja González complained that people searching for his name on the Internet would discover two brief newspaper accounts in January and February 1998 available on the site of the La Vanguardia newspaper. These stories were public announcements "mentioning Mr Costeja González's name" in connection "with attachment proceedings for the recovery of social security debts." Costeja González argued that the ability of the public to access these stories violated his rights under the European Data Privacy Directive, and he asked for the newspaper to delete his name and Google to remove links to the newspaper accounts.
>
> *(Balkin 2018, 1149)*

This is a ruling that has generated, however, a sort of "transatlantic divide": the United States of America, the jurisdiction where most of the important technology companies today have collected and processed data, did not welcome this judgment, since it has created a legal category far from their way of regulating technology in relation to the freedom of expression and the First Amendment. A legal category will moreover require rethinking part of their business model to meet the many requests for removal of information. In practice, this means that Europe is already prospecting sanctions that are often disproportionate to force large providers to react quickly to the many issues that are arising with regard to the need to remove contents (especially issues that have a major impact on public opinion).

As McNealy has stated, although EU member States hail the creation of this right to be forgotten as improving individual privacy rights, "such a right creates a problem for U.S. online news organizations. Not only does such a law come into direct conflict with protections found in the First Amendment, but it also conflicts with traditional privacy jurisprudence, which states that information made public cannot become private again" (McNealy 2012, 120). However, a technological regulation that is only intended to endure the threat of sanctions, and which does not arise from a process of dialogue and confrontation between Europe and the United States of America, is likely to prove very fragile in the short term.

At the same time, public opinion, political and jurist demands for the deletion of social network data is often raised in the wake of tragic events that make it clear that is impossible to interrupt the circulation of digital data after it has entered digital or social network platforms. The most striking case in 2016 in Italy was that involving Tiziana Cantone, a girl who committed suicide after having encountered the practical difficulty of removing sensitive information from the social network Facebook and from prominent web sites and search engines. This was not a case of the applicability of the right to oblivion in the Google Spain case mentioned earlier where the circulation of information was the result of a crime. Rather in this instance it concerned the resistance and technical persistence of online digital data

and the perennial memory capacity of the network. Often similar popular upheavals, however, lead to regulatory reactions that are sometimes characterized by rashness and legal inaccuracy, and which criminalize technology rather than focus on the heart of the problem. In many countries, for example, after tragic events related to the impossibility of removing data online, reactions have included eliminating the possibility of anonymity on networks, imposing more serious penalties for providers, increasing penalties related to the crime of defamation or revenge porn, increasing personal and companies' liability for online activities or forbidding encrypted communications between individuals.

We now find ourselves in a situation where memory cannot be removed or that its removal has become political. Are we heading towards a situation where memory will be conditioned by search engines, algorithms and persistent data, and where, above all, forgetfulness will no longer be possible and muddled memories will always be ready to be resurrected? Or, on the contrary, will legal remedies be able to guarantee everyone the right to be forgotten? Surely it will not be a political struggle fought against technology, but *with* technology, although data processing today is more about profiling the user and gaining profits from his actions than about protecting rights.

A new focus on forgetfulness will probably also require a radical change in the economic and profit maximizing focus of an information society. McNealy, citing Koops (2012), defines correctly from the literature three forms of the right to be forgotten: the right to have information deleted after a certain time, the right to have a "clean slate", and the right to be connected only to present information (McNealy 2012, 121). According to McNealy, "the first conception of the right centers on the idea that individuals should have the opportunity to require other individuals and organizations in possession of information about them to erase it", whether individuals or another person upload(s) information about the person online (McNealy 2012, 121). McNealy further notes that

> difficulties arise with enforcement of such a right because multiple parties exist that might be in possession of the personal information, as well as the possibility that some possessors of information might be required to retain information under the law.
>
> *(McNealy 2012, 121)*

The second and third forms of the right to be forgotten, the clean slate and the right to only be connected to current information, according to McNealy, "are similar. Both center on the idea that individuals can grow and change, and should not, therefore, be forever connected to information from the past that could be damaging [. . .] In these cases, individuals, for the most part, do not have the spectre of past ills or bad decisions available for others to use to judge them. The right to be forgotten would then allow people to "shape their own lives," instead of having the memories of others do so for them (McNealy 2012, 121–122). To this end, Murata and Orito offer this definition of the right: "An individual has the right

to be free from any use of information concerning him/her which causes harmful effects on him/her" (Murata and Orito 2011, 199). While the right to portability has already been recognized in legislation, particularly in the GDPR, and the Google Spain ruling, it poses far more complex social and political problems that arise in an information society.

The impossibility of a technological oblivion

In an information society without oblivion, people live constantly without the option of "undoing" an action (for example: permanently delete a post, or an e-mail message), or having a second chance. In a pre-digital era, Balkin notes, "old newspaper articles that contained embarrassing information were quite literally yesterday's news. People threw away the old copies and one had to go to library or some other location where the archives were stored" (Balkin 2018, 1203). In the digital age:

> organizations publish but do not delete. Instead, old articles are freely searchable in newspaper archives, which remain online. This fact changes the nature of a newspaper as an institution of the public sphere. The newspaper is no longer simply a report of the day's events, to be cast aside tomorrow and stored, if at all, in a relatively small number of libraries and other archival locations that are not quickly and easily accessible to the public. Instead, the newspaper becomes an increasingly important and valuable online archive. It becomes an institution of memory that is widely and easily accessible through search engines. Newspapers become important records of history experienced in real time that remain present for people to search and read days, months, and years later.
>
> *(Balkin 2018, 1203)*

When embarrassing material was published in newspapers in the past, Balkin concludes,

> the subjects eventually enjoyed practical obscurity when the newspapers were discarded (so that access to older newspaper stories was limited to those who visited libraries and archives). This practical obscurity has vanished because of search engines. Embarrassing articles may show up in search engine results and continue to appear indefinitely. By targeting search engine providers, the right to be forgotten attempts to restore the practical obscurity (and thus privacy protection) of the pre-digital era.
>
> *(Balkin 2018, 1203)*

Complete oblivion – intended as the real and concrete possibility of *erasing* or, better, of *destroying* any information that is no longer of public interest – is *incompatible* with the very architecture of an information society and the social networks and search engines

that make it up. Once information has been rendered in digital form and entered into the articulated set of connections that make up information society today – whether a blog, a WhatsApp chat, a social network discussion, or a comment – it is technically impossible to ensure its permanent removal and prevent it from returning online. Certainly, sophisticated operations can be carried out to mitigate the exposure or visibility of that information (such as the de-indexing, or the removal from an archive) to allow relative and momentary forgetting, but the process of digitization itself, and the ubiquity and capillarity of today's widespread digital technologies, makes it unrealistic to think that removing data from social networks after its digitization will secure its death. The risk that it will come back is always and consistently real.

Legal talk about oblivion in the digital world is, in many respects, thus a fiction. It promises the possibility that data can remain "under the surface", not become "trendy" or a subject of widespread attention. But can any more than that be expected? Today the most viewed content is that on top of the list of search engine results and the attention of the navigators typically never moves – unless there is a targeted specific interest – beyond the first two or three pages of results. All those operations that aim to tweak, mitigate, and reduce the visibility of certain information, may indeed have a practical effect and provide a good degree of protection of the rights of the subject. It thus appears that immediate forgetting can be controlled, but not the long-term oblivion. The latter possibility of being forgotten and erasing memory at the press of a button, recalls scenes such as in George Orwell's movie 1984, where Winston Smith works as a "censor of history" and removes articles from Times pages that are no longer politically acceptable. At the same time, some scholars have argued that we should worry less about how to *remember*, and more about how to *forget*. Viktor Mayer-Schönberger, in his book, *Delete*, has very carefully recalled, in many ways, how the inadequacy of human memory was, in reality, a valuable asset and could, over time, facilitate our social interactions (Mayer-Schönberger 2011). Past rumours, mistakes, gossip and conflicting thoughts were usually lost with the passing of time, giving people the ability to overcome imperfections in their lives and to start again, or simply leave behind parts of their past because they are no longer relevant.

Yet, as I have been arguing, this possibility of abandoning the story behind us has gradually been eroded by modern technologies. Mailboxes, social networks and online archives are perpetual extensions to our fallacious memories. Information is no longer lost even if we do not keep it in mind or in memory due to our human limitations. We can no longer move on from our past, and even small and irrelevant parts of it can return.

In some areas, for example criminal law, forgetting is not only considered acceptable and socially useful, but is considered *vital*. Annotations of small offenses are usually removed from official documents after a period of time, such as when small offenses committed by minors are cancelled before they become adults. The reason is simple: mistakes in judging youth and minor offenses must not prevent a person's reintegration into society. Ultimately, a second chance is offered to a minor, something that technology now potentially prevents.

In an information society that does not offer forgetting as a possibility, any behaviour that leaves a digital trace in the present can do immediate damage to a person's life but also have future damaging consequences. Often, immediate effects can be the least damaging: in social networks, certain events may attract attention for a few days, unless they concern public figures, and then remain quiescent. The potentially greater danger is that information will remain for the next ten or fifteen years or beyond and searchable with likely more and more sophisticated algorithms and search engines.

Conclusions: three forms of data oblivion

The information society is becoming increasingly complex. Algorithms, big data, the speed of information transmission and frequent violations of the privacy of the individual radically condition the life of digital data, and the ability to delete it and to guarantee a right to be forgotten. Legal, economic and political choices must be made together to address three forms of oblivion and their related interests and rights: i) social, iii) technical, and iii) legal oblivion.

Social oblivion concerns the persistence and circulation of personal data in information societies; the *technical* relates to the resistance of technology to the removal of data or the re-submitting of contents that were once deleted but are still present on personal computers and other devices and could be circulated again on the Internet; and *legal* oblivion refers to forgetting data, deleting information, and the de-indexing of news that has been elaborated by legal means through case law or norms. These three types of oblivion are inextricably linked and yet sometimes mutually incompatible: legal obligation can be guaranteed by a Court but may not become effective (and implemented) for technical reasons, or deleted data can continue to circulate in various social networks. The technological has, in fact, completely changed and challenged the legal form of oblivion.

Social oblivion is the right to be forgotten in a society, in the context in which one lives, with particular reference to certain aspects of one's life. It is the guarantee of being able to "start over again". However, social oblivion is now interwoven with technological oblivion, because people's information can be retrieved from the Internet, search engines and social networks, that make up an information society. If data is online, it will have the capacity to stay forever and vulnerable to being circulated thereby multiplying its effects.

Technically, forgetfulness is more similar to erasing and making digital data disappear. As previously noted, there is much scepticism about the possibility of a complete and persistent elimination. However, in many cases, technical forgetting can at least serve to make certain information less visible or more difficult to search. Legal obligation is, perhaps, the most definite (and restrictive) form of oblivion and is related to requirements elaborated by jurisprudence or norms. The connection to technological oblivion though often shows its limits. The right to be forgotten, according to the landmark Google Spain case, should now also be understood as the right to be forgotten by *search engines*: a protection granted to

all subjects who do not want to be remembered online for certain things they have done in the past, allowing the possibility of de-indexing that information from a search engine. However, the European judgment mentioned earlier refers to data that are not defamatory, false or confidential but true and originally and legitimately published but which, over time, have become unsuitable for to the more recent personal identity of the subject "portrait" and as such are no longer current or of public interest. Perhaps, this last point is the most delicate: how does this new right to forget deal with the right to chronicle and the right to publish everything that is related to facts of public interest? Should chronicle rights prevail over the right to personal identity, in the event that there is a public interest in the dissemination of facts, or not?

In theory, the right to oblivion (borrowing, this time, a typical United States approach) should not apply to information that has a public interest: for example, information related to personalities who hold public office, or political activities, and have a lower expectation of privacy. From this point of view, it seems to be a right strictly related to the concept of personal identity and to the representation of an individual in society. The Italian Privacy Authority has tried to clarify this point by stating, recently, that the right to be forgiven for legal action of particular gravity, and whose proceedings have been concluded, cannot be invoked since the public interest prevails to know similar news. The subject matter of the Authority's attention began in 2006 and ended in 2012, and concerned a former municipal councillor involved in an investigation. The person concerned claimed that, since he had no public office and was now in the private sector, the presence of news dating back to ten years before and now of no interest to him caused harm to his image, privacy and career. The Authority states that it is true that the time that has lapsed is the essential component of the right to oblivion, but also that if the information requested for removal is related to serious offenses and has caused a strong social concern, then deletions must be evaluated with less favour.

What this case highlights is that the relation between history, memory, traditional oblivion and technological forgetting and procedures necessary for the latter to be activated is not straightforward, and can create practical problems for both subjects and experts. What this reveals is that we are in the presence of a typical case where a "classic" theme has to be reinterpreted in a new light to adapt it to, in the case of an information society, the world of search engines and social networks. The judgment of "Google Spain" and the decisions of local independent, administrative and judicial authorities have obliged all industry operators to reason on certain aspects, and these nuances also affect citizens who want to exercise this right. In particular, the process of de-indexing now appears to be reserved in the first instance to search engines (which will make decisions based not only on the indications of the European Court but also on the basis of criteria and parameters established internally), secondly to independent authorities (for example the Data Protection Authority in Italy) or traditional courts, which will be able to operate on the basis of criteria different from those adopted by the search engines decision-makers.

It is clear, first of all, that there is a need for a case-by-case evaluation of each request as the right to be forgotten is closely linked to the subject's identity, personhood and civil rights. Technically, however, a case-by-case assessment can create enormous problems. After the European decision, many requests for removal have been made. According to the latest transparency reports of Google from May 2014 to September 2018 there were 2.7 million requests for de-indexing.[2] If requests continue to increase, providers will need to reorganize their business models such that human assessors can seriously analyse them on a case-by-case basis. Such a requirement has been implemented via Article 71 of the GDPR. It states that a person should have the right not to be subject to a decision (which may include a measure that takes into account personal aspects that concern them) based solely on automated processing and that has legal effects that significantly affect them. By preventing certain automated decisions without human intervention legal authorities have thereby foreclosed what could be one of the worst Kafkaesque nightmares.

Yet another consideration in this complex picture is the problem of time: the interpretation of a right to oblivion in an information society is linked to a conception of the passage of time. Data requested for removal must be "old" but, in the absence of certain parameters, how this will be observed will be subject to yet unknown interpretation practices. What is also important is the rule that oblivion, according to the "Google Spain" decision, should not apply to data related to serious crimes and sensitive social alarms or to facts that have raised great controversy, attention or public debate. The right to oblivion thus aims to preserve the publicity of data that concerns the individual and their life in a society, but which can put into danger their personal rights and reputation.

It is indeed appropriate to carefully assess the difference between the de-indexing of information – a category introduced by the European ruling – and the deletion of data. In fact, the right to forget in an information society may well become the de-indexing of data – that is to say that the data is not deleted but ceases to be indexed, that is made visible by search engines (for example, in the recent GDPR).

To help the user to get even better in this maze, Google stated in a report, via its Advisory Council, some parameters and rules of conduct in an effort to clarify this issue. These rules show the approach that Google has decided to take to address the requirements forced by the ruling and the interpretive criteria it deemed appropriate to adopt. The authors of this 2015 report, reunited in a "Advisory Council to Google on the Right to Be Forgotten", state, in the beginning, that:

> We were invited, as independent experts, to join the Advisory Council to Google on the Right to be Forgotten following the Court of Justice of the European Union's ruling in Google Spain and Inc. vs. Agencia Española de Protección de Datos (AEPD) and Mario Costeja Gonzalez C131/12 ("the Ruling") in May 2014. Google asked us to advise it on performing the balancing act between an individual's right to privacy and the public's interest in access to information. (p. 1) [. . .] We were convened to advise on criteria that Google should use in striking a balance, such as what role the

> data subject plays in public life, or whether the information is outdated or no longer relevant. We also considered the best process and inputs to Google's decision making, including input from the original publishers of information at issue, as potentially important aspects of the balancing exercise. (p. 2). The Ruling has been widely referred to as creating a "Right to be Forgotten." This reference is so generally understood that this Advisory Council was convened to advise on the implementation of this right. In fact, the Ruling does not establish a general Right to be Forgotten. [. . .] Implementation of the Ruling does not have the effect of "forgetting" information about a data subject. Instead, it requires Google to remove links returned in search results based on an individual's name when those results are "inadequate, irrelevant or no longer relevant, or excessive." Google is not required to remove those results if there is an overriding public interest in them "for particular reasons, such as the role played by the data subject in public life.
>
> *(p. 3)*

What is clear is that implementation of the Ruling does not have the effect of "forgetting" information about a data subject. Instead, it requires Google to remove links returned in search results based on an individual's name when those results are "inadequate, irrelevant or no longer relevant, or excessive." Google is not required to remove those results if there is an overriding public interest in them "for particular reasons, such as the role played by the data subject in public life" (p. 3).

In the report, first of all, four criteria (none of which are decisive or predominant) are highlighted, on the basis of which the company considers a correct assessment of the possibility of de-indexing certain data: i) the role of the subject concerned; ii) the type of information covered by the request; iii) the source of information, and iv) the passing of time. The first criterion, the role of the person concerned, is set out clear in the document: the role of the person concerned in the public context must be emphasized depending on whether it is in the order of: a) subjects with a clear and effective role in the public dimension, and the greater the public relevance, the smaller the chance that a de-indexing request will be accepted by Google in the light of the overriding public interest in seeking information; b) subjects without a prominent role in public life, and de-indexing requests should be easier to find by the search engine; c) subjects with a public limited role in specific areas, where is not possible to detect a greater or lesser likelihood of removal (pp. 7–8).

As regards the type of information that can be the subject of the case, the Advisory Council considered it appropriate to define two categories of "data type". The first category includes images or movies that represent the person concerned, information about his sexual life and economic situation, identification and authorization credentials, sensitive data, contacts, minor information, untruthful news or data that expose the person concerned to harm. In relation to these news and data, confidentiality is presumed to prevail, although the existence of a public interest might justify an exception to this principle.

In the second category, however, all the information normally held by a public interest would be covered. These include information on religious issues, political debates, or information related to public health and consumer protection or criminal activity, or data that contribute to debate on topics of general interest or with a historical interest, as well as information relating to scientific research or forms of artistic expression. In relation to this information, public interest in the news, and radical transparency (Lessig 2009), are presumed to prevail over a right to oblivion.

Referring to the source of the disputed news, Google claims that, in verifying the existence of a public interest, the source of the information and its purpose must be considered. For example, a major public interest can be considered to be based on news spread in journalistic activity or, in any case, on the activity of authoritative information sites. Finally, there is a temporal factor, it is said, that is essential in determining the existence of the right to de-index. Time is of particular importance in cases where the person's public role is actually changed with the passing of time.

When submitting requests for de-indexing, the interested party must provide all the relevant information so that the site managers can make an appropriate assessment. As for the relationship between site managers and the de-indexing process, the Advisory Council also recommends that the search engine operator, as a "good practice", must indicate the removal of the content from the website, within the limits provided for by law. As is clear, the margin of discretion is high, and the earlier-mentioned parameters provide first and useful guidelines, but leave room for wide interpretation that, on the one hand, could allow better tailoring of a decision to the specificities of a concrete case and, on the other hand, may protect the rights of individuals and respect freedom of information and free speech.

For Balkin, the right to be forgotten raises three issues (Balkin 2018). The first one is called "Collateral Censorship". In Balkin's opinion, "the right to be forgotten is a classic example of collateral censorship. Instead of going after the speaker, the state targets the infrastructure provider, and it threatens to hold the search engine company liable if it does not delink embarrassing articles from newspapers. The government puts pressure on the infrastructure owner to muffle (but not completely silence) the voice of the original speaker. The speaker is not completely silenced because if one knows the URL of the offending article, one can still access it; but of course, the point of the delisting is that without a search engine link most people will not be able to find it" (Balkin 2018, 1203).

The second issue is that the right to be forgotten "threatens the global Internet because the concern is that courts will eventually require global delinking as the appropriate remedy" (Balkin 2018, 1204). The last issue is that "the right to be forgotten is an example of how nation states (and in this case, the European Union) have tried to coopt private infrastructure owners and their capacities for private governance" (Balkin 2018, 1210).

What the right to be forgotten has reignited is the political issue of the right to oblivion and its impossibility in an information society where digital data are

susceptible to potentially infinite preservation and circulation. The protection of digital data is no longer about an alternative between deletion and non-deletion, but the choice between, at the very least, three different options: the total deletion from the source website, the de-indexing of the entire content and the de-indexing by name. The choice of one option rather than another involves different repercussions with regard to the balance between the rights of the author of the content or of the website of origin and those of the person concerned.

The social relevance of search engines as a means to access information on the Internet, together with the presence of monopolistic and oligopolistic positions, provide the conditions to control indexing and de-indexing. Regarding the right to de-indexization and, in a broader sense, the removal of search engine results, some political and technological choices could improve the situation. For example, the adoption of uniform procedures, guaranteed and transparent, to delimit, circumscribe and control the decision-making power and the discretion of the search engine, could limit dominant positions in the existing framework.

Addressing of all these issues and the interconnections between technological, social and political issues within this framework is proving to be very complex. The renunciation by the United States of part of the control by delegating to search engines and large platforms initial decision-making on the deletion of information, alongside the exponential spread of artificial intelligence, algorithms and big data, are pushing away the possibility of effective legal control. The provisions contained in the GDPR with reference to the portability of data and, above all, the more powerful right of deletion, are useful attempts at rebalancing rights, and at any rate, illustrate a complex site of data politics.

Notes

1 GWI Social – Globalwebindex's Quarterly Report on the latest trends in social networking, Q1 2015, www.thewebmate.com/wp-content/uploads/2015/05/GWI-Social-Report-Q1-2015.pdf, p. 5.
2 Avail. at: https://transparencyreport.google.com/eu-privacy/overview?hl=it

References

Balkin, J. M. 2018. "Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation." *University of California Davis Law Review* 51, 1149–1210.
Bauman, Z. 2000. *Liquid Modernity*. Cambridge: Polity Press.
Brock, G. 2016. *The Right to be Forgotten – Privacy and the Media in the Digital Age*. London: I. B. Tauris.
Carroll, B. and K. Landry. 2010. "Logging On and Letting Out: Using Online Social Networks to Grieve and to Mourn." *Bulletin of Science, Technology & Society* 30 (5): 341–349.
Carroll, E. and J. Romano. 2010. *Your digital afterlife: when Facebook, Flickr and Twitter are your estate, what's your legacy?* San Francisco: New Riders.
Floridi, L. 2011. "The informational nature of personal identity." *Minds and Machines* 21 (4): 549–566.
Floridi, L. 2012. "Big data and their epistemological challenge." *Philosophy & Technology* 25 (4): 435–437.

Floridi, L. 2014. *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford: Oxford University Press.

Garber, M. 2016. "Enter the Grief Police." *The Atlantic*, January 20. URL: www.theatlantic.com/entertainment/archive/2016/01/enter-the-grief-police/424746/

Herrera, T. 2017. "Is Your Digital Life Ready for Your Death?". *The New York Times*, January 18. URL: https://www.nytimes.com/2017/01/18/technology/is-your-digital-life-ready-for-your-death.html

Hildebrandt, M. and BJ Koops. 2010. "The challenges of ambient law and legal protection in the profiling era." *The Modern Law Review* 73 (3): 428–460.

Koops, B. J. 2012. "Forgetting footprints, shunning shadows: A critical analysis of the 'right to be forgotten' in big data practice." SCRIPTed 8: 229–236. DOI: 10.2966/scrip.080311.229

Lafrance, A. 2016. "Facebook Goes Full Nietzsche, Declares Users Dead." *The Atlantic*, November 11. URL: www.theatlantic.com/technology/archive/2016/11/facebook-digitally-murders-everybody-because-2016/507506/

Lessig, L. 2009. "Against Transparency." *New Republic*, October 9. URL: https://newrepublic.com/article/70097/against-transparency

McNealy, J. E. 2012. "The emerging conflict between newsworthiness and the right to be forgotten." *Northern Kentucky Law Review* 39 (2): 119–135.

Malone, M. S. 2013. *The Guardian of All Things: The Epic Story of Human Memory*. London: St. Martin's Griffin.

Mayer-Schönberger, V. 2011. *Delete: The Virtue of Forgetting in the Digital Age*. Princeton: Princeton University Press.

Mittelstadt, B. D., P. Allo, M. Taddeo, S. Watcher and L. Floridi. 2016. "The ethics of algorithms: Mapping the debate." *Big Data & Society*. July–December: 1–21.

Moreman, C. M. and A. D. Lewis (ed.). 2014. *Digital death: mortality and beyond in the online age*. Santa Barbara: Praeger.

Murata, K. and Y. Orito. 2011. "The Right to Forget/be Forgotten." *Ethics in Interdisciplinary and Intercultural Relations*: 192–199.

Parker, L. 2014. "How to become virtually immortal." *The New Yorker*, April 4. URL: www.newyorker.com/tech/elements/how-to-become-virtually-immortal

Rodotà, S. 2015. *Il diritto di avere diritti*. Rome: Laterza.

Schneier, B. 2008. *Schneier on Security*. Indianapolis: Wiley.

Siciliano, R. 2013. "How Do Your Digital Assets Compare?" URL: https://securingtomorrow.mcafee.com/consumer/family-safety/digital-assets/

Steinhart, E. C., Y Nagasawa and E. Wielenberg (ed.). 2014. *Your digital afterlives: computational theories of life after death*. London: Palgrave Macmillan.

Wachter, Sandra and Mittelstadt, Brent and Floridi, Luciano. 2016. Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*, December 28, 2017. URL: https://ssrn.com/abstract=2903469 or http://dx.doi.org/10.2139/ssrn.2903469

Weiser, S. 2016. "Should Prince's Tweet Be in a Museum." *The Atlantic*, July 5. URL: https://www.theatlantic.com/technology/archive/2016/07/should-princes-tweets-be-in-a-museum/489776/

Wilmot, C. 2016. "The Space Between Mourning and Grief." *The Atlantic*, June 8. Avail. at: https://www.theatlantic.com/entertainment/archive/2016/06/internet-grief/485864/

Wong, C. 2013. "Can Bruce Willis Leave His iTunes Collection to His Children? Inheritability of Digital Media in the Face of EULAs." *Santa Clara High Technology Law Journal* 29 (4): 703–761. URL: http://digitalcommons.law.scu.edu/chtlj/vol29/iss4/5

13

DATA CITIZENS

How to reinvent rights

Jennifer Gabrys

Introduction

Air pollution is increasingly recognized as one of the biggest public health crises on the planet (Das and Horton, 2017). Indeed, next to climate change, air pollution is a major environmental event that is affecting cities worldwide. From Beijing to London, and from Tehran to Warsaw, cities are experiencing varying levels of air pollution that harm urban dwellers' health and that have significant economic costs. While world cities and manufacturing cities, resource cities and emerging cities all suffer from poor air quality, pollution levels are often quite disparate across these multiple sites. In New Delhi, instruments recently topped out at "999"[1] and were not able to register further increases in pollution levels. The environmental crisis of air pollution overwhelmed the devices that were meant to measure and, by extension, the data used to govern the ill effects of air pollution. Responses to air pollution across cities worldwide are now formed through a complex mix of expert-based monitoring networks, official air quality indices, public health guidelines, home filtration systems, breathing technologies, low-emission transport routes, citizen monitoring and political protest, along with local dynamics in the spread and concentration of pollutants, and international and intercontinental movements of air.

The "right to clean air" is variously upheld and observed in different urban environments. Although the official infrastructures for monitoring air pollution are meant to assure urban dwellers that constant monitoring, control and even care is given to the air that they breathe, ruptures in the systems and technologies of governance regularly occur. The expert practices and techniques that would ensure that urban air is breathable become the target for questioning and frustration. Urban dwellers at times doubt the accuracy of the air quality data that is made public; or they rail against the inertia within urban and national governments that they feel do little to improve air quality. For many urban dwellers, expertise folds in on itself in

these scenarios, and seems to become more of a defensive structure for "elites," and less of a process of accountability. Expertise alone, and expertise in its usual static configuration, is inexorably remade through the urban-environmental demands of air pollution. Citizen monitoring of air quality, and the citizen data that it generates, then become ways to challenge, contest and account for harmful environmental conditions and to pursue the right to clean air.

While data is often seen to be something that is collected about citizens (typically by large technology companies), there are just as many instances now arising of citizens generating their own data. Whether to document lived experiences through social media platforms, to sense air pollution to challenge governmental readings, or to address conflict in areas of development, citizens are collecting, analyzing and communicating data in order to operationalize new types of evidence. This chapter documents how citizen practices of using low-cost and digital sensor technologies to monitor air quality and changing urban environments in Southeast London generate distinct modes of citizen data, as well as specific formations of data citizens and data relations. In collaboration with the Citizen Sense research project, residents, workers and volunteers in the Deptford and New Cross neighborhoods of Southeast London took up air pollution sensing technologies to monitor air quality. Citizen sensing technologies were often located adjacent to construction sites and traffic corridors where rapid urban development was underway in order to demonstrate problems with air quality. Citizen generated data then became one way to attempt to intervene in and reshape processes of urbanization, especially as they contribute to environmental harm.

By discussing this specific set of citizen data practices tuned to urban environmental change, I investigate the ways in which data becomes a medium for democratic engagement, and how participants become "data citizens" through the collection and operationalization of data. I ask: How do multiple and different types of data and data practices, including citizen sensing, constitute distinct data citizens? In what ways do data citizens mobilize rights—to data, to the air, to the city and to political life? And how does citizen data potentially both supplant and reinvent rights as the way in which urban inhabitants mobilize in support of the urban realm? I take up these questions to consider what sorts of political subjects concretize through the collection and production of data to document, contest and argue for urban environments.

Above and beyond a right to clean air, citizen data practices on the one hand could become a way to materialize a distinct "right to the city." Drawing on the work of Henri Lefebvre (1996), this chapter examines how data citizens might express a right to the city through rights to data. For Lefebvre, the right to the city was a call to consider how participation in the city might be enabled for often-excluded urban inhabitants. Such participation spanned from the right to public space, difference, housing, political engagement, social life and even information technology. The right to the city complemented and potentially extended the "abstract rights" of citizens (1996, 157). Lefebvre was reinventing rights and their practice—as open ended and in the making—through his call to the right to the city.

As I suggest here, the right to data could be read as one articulation of the right to the city, where citizens might participate in the analysis and making of evidence in order to support and create urban life. This is a particular way of understanding the right to the city through the right of citizens to generate, collect, analyze and communicate data that can dispute and question official accounts of problems such as air quality in relation to urban processes. Through these data practices, distinctive modes of data citizens that claim a right to the city could materialize.

However, the right to the city, despite its call to accommodate diverse urban inhabitants, could on another level compel particular modes of participation in order for urban citizens to be recognized as such—as *active* participants, made visible through recognizable rights claims or legible interventions. In relation to the citizen sensing of air quality in Southeast London, I then consider how practices of participation, rights and citizenship do not always unfold so easily. Drawing on Lauren Berlant (2011) and her discussion of proliferating, multiple, intimate and even depressive modes of citizenship, this chapter seeks to differentiate modes of data citizenship and urban citizenship that challenge—and potentially reinvent—the right to the city and the right to data through alternative approaches to participation and citizenship. Data citizens might in one way be less oriented toward the overt ambitions of rights, and more engaged with finding provisional techniques for staving off and surviving dispossession, pollution and injustice that often accompany increasing urbanization. A right to the city promises powers of engagement and transformation that can seem to be remote possibilities for many urban dwellers. For urban citizens turned data citizens, rights (even of the Lefebvrian sort) could appear as a murky form of "cruel optimism" (Berlant, 2011).

This chapter then considers how the mobilization of data and evidence can potentially displace and reinvent rights—including the right to the city. Rather than refer to rights, here data—"the facts"—becomes a moveable baseline for making arguments in support of urban life, but without a clear arrangement of rights as such. Rather than appeal to a right to housing, for instance, urban inhabitants might demonstrate the amount of new construction that is not affordable to local residents, thereby creating data-based arguments that generate particular observations about unlivable conditions. In contrast to citizens who practice a right to the city, here data becomes a stopgap measure to sustain an urban way of life that is continually under threat, but for which very few rights exist. Data citizens form where the right to the city meets cruel optimism.

Citizen data could, in this sense, be a technique that manifests on the other edge of cruel optimism. People who might not feel that rights are a clear point of political attachment, in principle or practice, instead create evidentiary techniques to challenge the dispossession, environmental damage and injustice of neoliberal urbanization. Data for Black Lives is an example of such a movement that involves developing alternative data collection and analysis techniques to create new narratives and to demonstrate systemic racism. In this practice, rights are not always self evident, since there are many rights that black peoples have but that are often not protected or observed, and there are many more data-based

arguments that might be made that do not have a clear reference to rights. Instead of data mobilized to support rights as such, data could then be mobilized to support struggles for survival in the absence of rights. Data citizens are not identifiable here through the usual categories of membership, whether to nation-state or group, or through reference to designated rights as such, but rather through operational practices that form particular political subjects, relations and communities by working with and through evidence. This is a way of reinventing rights in practice by working through concrete struggles to evidence harm and to generate more livable urban worlds. I now turn to consider how urban citizen sensing data is mobilized in these ways.

Right to the city, right to data

In a chapter on data, it might seem circuitous to engage first in a discussion of the right to the city. But such an engagement is deliberate, since it brings about another approach for investigating the distinct contributions to be made by urban citizen sensor data on environmental conditions such as air quality. Data might often be an expression of certain types of rights when connected to online and social media spaces, such as rights to privacy, to be forgotten, to data protection, and to open data. Here, I take up another way of thinking about how data citizens might form through the right to the city. Lefebvre developed the "right to the city" as a way to move from "thought to action," and to break open the abstract—and expert—urban science discourses that often seemed to operate on a self-referential and "meta" level (1996, 152). The "right to the city" is first and foremost a statement about how participation in the city makes and remakes the city—as an "*oeuvre*" (1996, 173–174). As a place of exchanges and encounters, the city has the potential to spark into being new urban societies, new urban subjects and new praxis (1996, 149–150). The right to the city, as well as the right to difference, the right to ways of life, the right to inhabitation and the right to information, are thoughts that galvanize such praxis (1996). Indeed, for Lefebvre, "a thought which tends towards an opening leads the struggle" (1996, 63). The praxis that Lefebvre imagines and proposes is not one of generating manageable outputs. Instead, it is a praxis that seeks to create the city as an ongoing collective project.

Praxis then involves the actual undertaking of the right to the city. Praxis is a constellation that forms through the exploration of the right to the city as a thought and struggle. Because it constitutes "social life," for Lefebvre praxis can be investigated sociologically, since "sociological thought seeks an understanding and reconstitution of the integrative capacities of the urban as well as the conditions of practical participation" (1996, 153; cf. Balibar, Cassin and Laugier, 2014). As a way of working across formative principles and concrete ways of life, praxis also demonstrates how "there is an urgent need to change intellectual approaches and tools" that might incorporate new and unfamiliar methods (1996, 151). This research on citizen sensing in many ways forms as an inquiry into such conditions of practical participation, while it also seeks to experiment with new and unfamiliar methods

of urban inhabitation. Praxis captures the thinking and the doing, as well as the generation of new thoughts and practices that can transform customary approaches, here to urban environmental problems.

The "right to the city" is a concept that has been extensively discussed and debated in urban research and practice. While there is not space here to summarize this vast field of research, there are two particular strands that are particularly relevant for this study. The first is an emerging area that looks at Lefebvre's under-examined reference to the right to information as a key component of the right to the city. Indeed, Lefebvre (1996) suggested that the right to information would go along with the right to the city and the right to difference so that urban inhabitants (or *citadins*) could provide their own accounts of urban life. Here, Joe Shaw and Mark Graham (2017) have discussed how citizens might have the right to information that is open, transparent and in support of urban democratic life. They contrast this project with the less generous practices of Google, which they suggest constructs exclusionary cities through its search and mapping functions. The second area of research is a more long-standing set of investigations that engage with the right to the city as a way to support and understand grassroots, DIY and citizen-led projects. Writers from Margaret Crawford (2011) to Mark Purcell (2002) and Don Mitchell (2003), as well as Doina Petrescu and Kim Trogal (2017) have elaborated on the ways in which cities form through concrete actions and engagements.

These studies are important reference points since on the one hand they demonstrate that the right to the city can both mobilize and capture collective urban energies, but on the other hand that it also leaves open many questions about who has the right to the city, and under which circumstances. While the right to the city could seem to be characterized by a set of universal requirements such as housing, there are always more (incomplete) rights emerging. The practice of realizing these rights is also in process, and often marked by struggle. The rights that materialize through practice are less a formal legal script, and more a dynamic and experimental opening into urban life. They are "rights in the making" (Lefebvre 1996, 179). In this way, the undertaking of the right to the city is more relational rather than teleological, since it is less focused on arriving at a finished urban form, and more attuned to the ways of life that are experienced and sustained, as well as the political subjects that urban inhabitants become in these collective urban projects (see also Harvey, 2008).

For Lefebvre, rights are aspirational and pursued as part of a hopeful praxis. The right to the city is a proposition for how cities might be made more expansive, generous, creative and accommodating for all urban citizens. The right to the city then sparks another way of thinking about rights and the city by forming a thought about collective urban life that galvanizes everyday practice. The right to the city here goes beyond human rights or liberal rights, and to a certain extent encompasses a sort of "new right" similar to environmental rights (Isin and Ruppert 2015, 23). But in another way, the right to the city involves the reshaping of rights as something collectively constituted, something made in everyday practice, and

something that forms the urban worlds that we inhabit. So too would the right to information, or here the right to data, involve a practice of rights in the making that seeks to realize ways of engaging with data that contributes to the formation of political subjects, relations and communities.

As compelling as the proposal is for the right to the city, in the contemporary post-1968 context it is difficult not to read Lefebvre's text alongside other articulations of rights and citizenship that also grapple with utopic or optimistic political impulses, only to point out that there may be a harsh aspect to such political aspirations. Berlant (2011) has engaged extensively with the promises and perils of citizenship, along with modes of democratic engagement that are affectively binding, yet which also can be productive of "cruel optimism" in their brandishing of possibility that does not manifest. While right to the city projects have achieved significant presence through analyses of grassroots urbanism (Mitchell 2003; Petrescu and Trogal 2017), and have also been integrated into urban campaigns and United Nations Habitat goals (UN 2017), still for urban dwellers in many parts of the world the right to the city is ever-elusive in how it might materialize in lived urban experience that is marked by ongoing struggle.

The right to the city seems to promise that a more democratic and livable city could be realized through praxis. Yet in the process of attempting to realize collective urbanization, many struggling urbanites are often worn out and worn down. As Berlant writes about such political attachments, people are "worn out by the promises that they have attached to in this world" (2011, 28). Moreover, for all of its evocative power, the right to the city can seem to work within a universal, normative, masculine, and actively enabled form of urban citizenship. Such citizenship would in part require that urban dwellers struggle and confront the city in its injustices and exclusions, often in public forums and settings. While there have been a number of studies focused on the tactical, subversive and DIY approaches to urbanism (e.g., Iveson 2013), these are often temporary interventions that are undertaken by particular groups of urban dwellers, and which do not necessarily extend to remaking processes of urbanization, or to challenging the "liberal monohumanist premises" (Wynter and McKittrick 2015, 11) of urban citizenship.

While the right to the city is meant to be equally available to all, these rights are conjectured through various modes of being present and struggling for those rights. Indeed, it could be said that for many people, the very act of attempting to take up the right to the city would attract considerable harm and violence, since people of color, women, disabled people, migrants, and many other urban dwellers are not as easily entitled to urban practices that implement a right to the city. In other words, they do not necessarily have "the right to have rights" (Arendt 1951; DeGooyer et al. 2018). Indeed, as Purcell has noted, "The right to the city is not inherently liberatory" (2002, 103). Not all urban inhabitants will be equally empowered, if at all, in the struggle for the right to the city. Those who do realize empowerment could create new exclusions for other urban inhabitants. Although Lefebvre sought to be "radically inclusive" where urban rights

extended to everyone and "not just to officially designated 'citizens'" (Crawford 2011, 35), urban struggles inevitably materialize in concrete ways that have different consequences for diverse urban inhabitants in how they are addressed.

While for Lefebvre the right to the city should be available to all—and he was interested in this being a diverse all—Berlant somewhat differently seeks a proliferation of citizenship, as well as the forms of relation that accompany these modes of citizenship. Rather than bundle rights into a practice available to a singular if diverse all, she tunes into the plurality of political subjects and the struggles they encounter. In this way, struggle is no less crucial for Berlant, but it is through struggle that even more modes of citizenship might be realized (Berlant and Seitz 2013). Engaging with this proliferation of modes of citizenship, I would suggest that such diversity of different political subjects also extends to a diversity of urban worlds inhabited. Citizenship is a sited, collective and relational practice that takes up and responds to a city in different ways. As Berlant notes, "Citizenship is the practical site of a theoretical existence, in that it allows for the reproduction of a variety of kinds of law in everyday life" (2007, 38). These modes of practice demonstrate commitments to struggle for worlds that might be more livable, but they also are unevenly available and show that failure is likely.

Failure, however, is not the flip side of success, but rather is a recognition of the pitfalls in praxis, where a thought that leads a struggle will also encounter the gritty conditions in which struggle unfolds. Failure in this sense involves the "impasses of the political" (2011, 4) where the usual modes of engagement become untenable. The reinvention of citizenship, rights, communities and the worlds that are made and sustained through political relations can, in these moments of impasse, begin to appear more viable. As Berlant writes,

> It may be a relation of cruel optimism, when, despite an awareness that the normative political sphere appears as a shrunken, broken, or distant place of activity among elites, members of the body politic return periodically to its recommitment ceremony and scenes.
>
> *(2011, 227)*

Such recommitment can involve paying attention to how political formations hold together, how they fall apart, and how they might be remade toward a "more livable and intimate sociality" (2011, 227). Striking an atmospheric note, Berlant suggests that "ambient citizenship," where the affective infrastructures of everyday life are tuned into, presents a way to move beyond the normative structures of governance to reveal how these more livable conditions might be realized (2011, 230–231). This is another zone of the political, which develops along with citizenship practices as part of the infrastructural commoning of political life (Berlant 2016).

Here I consider how "cruel optimism" accompanies the right to the city. The point of such an endeavor is less to lambast the aspirations of the right to the city, and more to engage with the inevitable complications and complexities that come from trying to engage in urban democratic processes. The right to data, moreover,

becomes one particular way in which the right to the city is pursued and expressed, and yet also can be derailed, whether through sclerotic urban governance structures, rigid arenas of expertise, or exclusionary processes for making an account of urban worlds beyond economic growth. The right to data, which is neither self-evident nor given, is then tied to ways of constituting distinct modes of data citizens that proliferate in these sites of urban struggle.

Creating data citizens

"Data citizen" is a term that is in broad use across industry, research and activism to variously describe the ways in which subjects are constituted as techno-political actors through their data practices. Within the tech industry, data citizen is used in a general way to suggest an ease and accessibility of participation with data technology, and with data analysis techniques more specifically. "Citizen" is often appended to digital technology to give a seemingly democratic gleam to these developments, from citizen sensing to data citizens. While "data citizen" is variously deployed to refer to the intersection of data and subjects, it can often be somewhat unclear what makes these citizens political subjects as such, and how data contributes to this formation. For instance, in what ways does political participation unfold through engagement with data? Are these practices direct and deliberate engagements to collect data to act on a political problem, or do they involve becoming entangled within infrastructures of data collection and mobilization that are necessarily political—or both?

Data citizens, as science and technology researchers Judith Gregory and Geoffrey Bowker discuss, can assemble through particular quantitative techniques such as those facilitated by wearable technologies. In their estimation, data citizens are constituted with and through "an ecology of microdata," rather than preceding this relation as such. "We are," they suggest, "*constituted* differently as data citizens at different technological moments" (2016, 220). Data citizens do not assemble with wearables through a premeditated plan to participate, but rather through the ecologies that are joined up through their ongoing data collection. Indeed, such data citizens might find that their "rights" to data are restricted if they attempt to access and use their own data or the data of others in these ecologies. Data citizens, in this sense, are not necessarily always working in a deliberative or democratic vein. If we consider the "conditions of practical participation" discussed earlier, there are multiple instances of participation that do not lead to a "right to" anything as such. Instead, participation can become the basis for further de-democratization, even while the term citizen is mobilized to suggest otherwise. Participation can in this way generate conditions of cruel optimism.

There are then multiple and competing modes of political subjectivity that materialize through the language and practices of "data citizens." These plural ontologies, far from consolidating into a single trajectory of empowerment, instead demonstrate the multiple uses and abuses to which data citizens can be put. Given that data citizens materialize through concrete practices, the modes and forms of

data citizens are inevitably multiple. The practices of gathering and mobilizing data can be crucial to the formation of data citizens as particular political subjects. Numerous writers, from Isin and Ruppert (2015) to Gabrys, Pritchard and Barratt (2016), have suggested that citizenship can be expressed and materialized through data practices, which constitute rather than predetermine political subjects.

In this way, data citizen as a term takes on distinct meaning in relation to citizen sensing. Here, data citizen as a term transforms and generates new perspectives on data and citizens through urban sensing experiments, where the creation of citizen data in relation to the urban environment can have particular effects. Environmental sensors are meant to enable and activate particular forms of environmental citizenship. They embody a version of the good life that is meant to be in reach for anyone, anywhere. Plugging in, activating a digital kit and joining a disparate community of users, are the steps to be followed that in principle should mobilize environmental rights in the making. Yet rather than unfold a more straightforward form of political engagement, citizen sensing kits and the citizen data they generate can often given rise to even more complex struggles with urban environmental life. Considerable work goes in to collecting and analyzing data sets, yet citizen data is often treated with suspicion and disregarded by regulators and industry. Indeed, even the right of citizens to monitor environments can be thrown into question, with practices, protocols and devices all subject to legal intervention and scrutiny (Kravets 2017; Pidot 2015). The promise of the political subject who is meant to form through the collection and communication of environmental data is then troubled. It is at this site of struggle that other forms of data citizens and urban citizens proliferate, less as fully formed legal actors, and more as persons attached to, yet haunted by, the promises of democratic life.

As one of the stated areas of inquiry for this collection, in order to understand and influence formations of data politics it is also necessary to understand how data citizens materialize (Ruppert, Isin and Bigo 2017). As Gregory and Bowker's work suggests, many studies on data citizens focus on technologies and data generated through wearables or social media, which perform particular expressions of consumer-subjects. In this analysis of citizen sensing, however, data citizens materialize through the production of citizen data using low-cost digital environmental monitoring technologies. Citizen sensing technologies are promoted as a way to encourage participation in environmental problems. Yet the process of sensing environments, collecting data, documenting and addressing environmental harm is also the site of an uneven formation of a citizen-sensing political subject. While citizen sensing technologies could very well reinforce and reinscribe these modalities of consumer-subjects, I suggest there are different ways of engaging with the possibilities of citizen data in relation to urban change and conflict that can rework both data citizens and processes of urbanization. The right to the city and the practices that it mobilizes become ways to rethink and rework the ongoing struggles with data in situ, as struggles that remake cities and their inhabitants, as well as their rights. If data practices contribute to the formation of citizens as political subjects, then it would seem they are also fused with the articulation of rights in the making.

This analysis examines engagements with data citizens and citizen data to consider how data is on the one hand produced in and connected to urban environments through sensors that monitor air quality; and on the other hand to study how citizens form environmental evidence that relates to their lived experiences. Urban sensing and the data it generates can become a way to make claims to and about urban environments by articulating individual and collective grievances about pollution, development, displacement and dispossession. Data practices in this milieu become expressions of urban citizenship, where a right to the city could be undertaken through sensor data and on-the-ground observations. In these re-articulations of rights to data and to the city, there are also surfacings of distinct modes of citizens and citizenship, as processes, relations, and communities, rather than fixed prescriptions for political life. Data citizens are constituted through practices of forming and making evidence, whether by generating their own data or compiling and analyzing diverse datasets—or both. But the data citizen as political subject is neither settled in relation to articulated rights nor easily identifiable through membership to a group such as a nation-state. Rather, data citizens can form through struggles with the erosion or absence of rights, and through the inability or futility of appealing to rights. Evidentiary techniques then become an ongoing process for materializing data citizenship as an affective and collective infrastructure for engaging with and intervening in urban worlds.

Citizen data, urban worlds

For the past six years, I have been leading the Citizen Sense research project, which investigates the use of DIY and low-cost digital environmental sensors. The research group has looked at existing sensor practices used to monitor radiation after Fukushima, and practices to monitor methane emissions at lost, abandoned and orphaned oil and gas wells in Pennsylvania, among many other forms of citizen-based environmental monitoring. One component of these investigations has been to look at the rise of digital sensors as environmental tools that are more readily available, and that are meant to expand the possible participants in environmental monitoring beyond expert scientists and technologists. This is the promise of sensors, but there is relatively little research to see how these practices actually play out.

Adopting a participatory and practice-based set of research methods, the Citizen Sense research group has identified communities and practitioners already using different sorts of environmental monitors, and then through a process of dialogue and distributed creation developed monitoring kits that could be used in response to environmental problems. While our research group assembles monitoring toolkits for adaptation and use, the process of building a monitoring infrastructure takes place with communities and in response to their specific concerns. In our most recent research spanning nearly two years from 2016 to 2017, we collaborated with residents of the Deptford and New Cross neighborhoods in Southeast London to monitor air quality in relation to traffic, development and industry emissions. This site of former industry, dockyards and a historic naval shipyard has undergone

successive waves of regeneration. Since at least the 1990s to the present day, communities have organized in order to be able to contribute to, respond to or contest processes of development. In the current context, the urban fabric in this location has been reworked and gentrified through new development schemes, master plans and public-private initiatives. Here, the city is being made and remade, less as an expression of the right to the city, and more as a set of developer projects that lead to ongoing contestations over the urban environment.

Indeed, a number of development sites were and continue to be actively contested by residents, where planning permission was sought (and in some instances, granted) to develop relatively unaffordable housing in the place of community gardens and social housing. Participants in this neighborhood have then been especially concerned about rapid rates of construction underway, as well as large (and typically luxury) housing developments yet to come that would significantly alter the area. One small area, Creekside, located on the eastern edge of Deptford had at least five separate development sites underway during the time of this monitoring study. Residents suspected that such developments were likely contributors to increased air pollutants throughout the development lifecycle. From demolition and site clearance, to construction and heavy goods vehicles, as well as increased density and traffic once development is completed, the environmental effects of construction can be felt for years. At the same time, the impacts of construction are inevitably bound up with the perceived lack of economic and social justice related to new developments, as people are displaced from social housing and often not able to afford to live in the area once the brunt of negative environmental effects from development had been endured.

In order to contest development, as well as to seek compensation from developers in the form of community development funds, many residents and community groups had undertaken environmental monitoring projects in order to demonstrate the ill effects of living with ongoing construction. From traffic counts to air quality studies using diffusion tubes, local citizens generated multiple forms of data about their environments. People also encountered, analyzed, and used data from governmental entities and industry, including in the form of planning documents in online portals; community meeting minutes; environmental impact and environmental assessment reports; official air quality data; construction company self-reporting on pollutant levels (including air, noise and light); utility company data on pollutants from national infrastructure projects (including air and noise); tree map data designating tree locations and numbers; tree removal applications; social statistics on population, density, and income; social media data (including Twitter and Facebook); crowdfunding data; petition data; word-of-mouth data (often on new development schemes); and many more types of data on the London Data Store and the Lewisham Borough website.

In these numerous engagements with data, environments and governance, people became data citizens in part through wrestling with these multiple forms of data. They analyzed data that was publically available, they sought data through FOIs, they documented events and environmental disturbances by creating their own

datasets, and they communicated and contested changes to the urban environment through these multiple data sources. Citizens also produced their own data often as a way to counter or qualify government statements and industry claims. People produced data in the absence of official monitoring networks, or where austerity measures meant that data was not sufficiently analyzed or acted upon. These multiple data practices are all ways of making evidence, in part through creating new citizen data, and in part through linking different data sources in new ways to create particular accounts of processes of urbanization, and to intervene in these processes. It was in this context that the Citizen Sense research group collaborated with residents to develop a citizen-led air-quality monitoring network that would generate data to be integrated into these complex and multiple data practices, and to research the ways in which data citizenship might materialize or transform.

Setting up a citizen monitoring network

Because many inhabitants in this part of Southeast London already had established data practices of various sorts, whether in the form of environmental monitoring or analyzing government datasets, our collaboration with communities then involved learning more about their data practices, while also engaging in dialogue, workshops, walks, meetings, and site visits to communicate about the particular configurations of citizen sensing technologies. Far from acting as "experts" with a singular way of accounting for urban environments, we contributed data practices that joined up with existing community infrastructures, while also investigating how these infrastructures could adapt and grow. We were, in the process, also becoming particular data citizens as we collaborated with inhabitants and learned more about concerns in the area.

Along with learning more about ongoing community campaigns in the area, we worked with residents to develop an air quality monitoring kit. In response to the area and concerns, we developed bespoke sensors that we called the "Dustbox." This device was a small plug-and-play sensor measuring particulate matter 2.5 ($PM_{2.5}$), a particularly hazardous air pollutant. The Dustboxes were housed in 3D-printed ceramic forms based on the shape of particulates when viewed under an electron microscope. The sensors were developed to be a tactical and affective device that would build up into an engaging community-monitoring infrastructure. We first held a workshop and walk to introduce participants to the device, as well as to discuss key monitoring locations and to address air quality topics. We then set up a monitoring network that included up to 30 Dustbox sensors monitoring $PM_{2.5}$ over a span of nearly 10 months, although the number of Dustboxes running varied throughout the monitoring period. Numerous visits were made to monitoring sites to set up devices, connect them to Wi-Fi networks, find suitable outdoor space for monitoring urban air, and troubleshoot along the way as devices went offline or required repairs (see also Houston, Gabrys and Pritchard, 2019).

In addition to setting up the Dustbox sensors, we built a database and online platform that we called "Airsift," where participants could view and analyze their data

in relative real-time. The Airsift tool was an attempt to investigate not just citizen sensing technologies as DIY devices, but also to work toward DIY data analysis as a key part of how citizen data could be investigated and used. Using Airsift, we found that this spatially dense community network of sensors allowed us to zero in on particular urban patterns, processes and distributions of pollutants. Often working at the scale of 1-hour and 24-hour mean levels of particulates, we could attend to the specific and comparative timing and distribution of pollutants in the area, which allowed us to gain a much more detailed picture of urban activities underway.

If citizens collect data but their data is closed down or inaccessible to analysis, then this practice might more accurately be referred to as crowdsourcing, since the data is owned and mined by actors other than the citizens who collect the data. Working with the community and building on our previous citizen sensing projects, we developed and refined the Airsift DIY data platform so that citizens could review and analyze their own data as well as other data in the network. We held workshops to discuss ways of analyzing and using data in support of community projects. Data in this sense was more than "open," since it was not simply a CSV file made available by a government entity in a data repository, for instance, but instead was embedded in the situated monitoring and data collection practices, as well as available for analysis, and mobilized within projects to advocate for the urban environment.

Working with citizen data and assembling data stories

As the monitoring network was forming as set of technology installations and urban relations, we worked with participants in a series of data workshops to introduce the Airsift tool, to work through analyses of different citizen datasets, to compare different monitoring sites, and to strategize about where else in the area might be useful to place monitors and gather data. The data that was accumulating from 30 Dustboxes sited across Southeast London began to inform particular modes of data citizenship. When arguments about urban housing were not heeded by the local council—as there was neither a specific "right" to be claimed, nor did people feel as though a rights claim would be respected—participants combined further data about air quality to form evidence about the impacts from ongoing construction.

As a register of urban environmental processes, the Dustbox citizen data began to unfold in relation to everyday urban life. Moments when air pollutants were registering particularly high levels became an event where participants would pool collective knowledge about industry activity, fires, high pollution drifting in from Europe or other events such as intensive construction activity that might help to explain peak readings. In this sense, quantitative sensor data did not provide an absolute or definitive reading of urban events. Instead, citizen data became most illuminated when multiple observations and other forms of data came together to corroborate and also transform lived urban experience.

We captured our collective findings from the 10 months of Dustbox monitoring in seven *Deptford Data Stories* that document and analyze the citizen data, as well as interweave the numerical measurements with on-the-ground observations and

images that form urban air pollution narratives. We discovered that major traffic intersections and construction activity, as well as the River Thames, all show up as likely pollution sources, often at levels well above the WHO 24-hour guideline of 25 μg/m^3. We also found that green spaces and sheltered gardens can have much lower levels of $PM_{2.5}$. The *Deptford Data Stories* provided a way to engage with citizen data beyond presenting measurements to instead link up lived urban experience with interpretations of the citizen datasets and proposals for concrete action in the urban realm. Citizen data in this case was not seeking to fulfill a regulatory function, but rather it was asking different questions and providing different insights about air quality pollution in relation to broader urban environmental processes.

Here, the right to the city and the right to data concretized into particular modes of data citizenship that drew on existing urban relations while building new ones, that tied into community projects while advancing other propositions for how urbanization might unfold. Processes of making evidence—the right to data—also created particular ways of claiming the right to the city. Yet these rights were unevenly recognized by local and national government, by industry and developers, and by other "stakeholders" who might respond to data citizens as they claim a right to the city, a right to data, and the right to clean air. Proposals made in relation to data gathered were then frequently developed in response to the impasses experienced and anticipated while advocating for the urban environment.

Plural data, plural urbanisms

While data citizens form through the multiple registers of urban environmental data, they also have the potential to challenge the usual ways of documenting and addressing environmental conditions. The practices of data citizens, furthermore, raise distinctly different perspectives on urban conditions. The right to data in many ways materializes not the right to *the* city, but rather the multiple cities that urban inhabitants traverse and bring into being. Indeed, one air quality officer I have spoken to about air pollution levels in London stated that there was little that could be done about $PM_{2.5}$ levels in their borough, as the annual average of 19 μg/m^3 varied by only +/-1 μg/m^3 across their monitoring area, and particulate levels were seen to be attributable to pollution traveling from outside of the immediate area, or even from Europe or farther afield. From the expert's-eye view it might seem sensible to agree with the intractability of this problem, even though annual $PM_{2.5}$ levels of 19 μg/m^3 are nearly twice the World Health Organisation (WHO) annual guideline of 10 μg/m^3. Yet expert practices and infrastructures are here attending to the problem of air pollution in a particular way, assessing data sets according to annual averages as a measure of compliance (or not) with air quality objectives. The numbers, which apparently capture the facts of air pollution in London, will not budge, and so it seems we are stuck with the air we've got.

But data citizens can offer a different picture of urban air pollution, where differently granulated patterns arise and distinct city processes come into view. Inevitably, in the process of researching and using sensor technologies multiple questions arise

as to the accuracy of devices, the actors who are able to put forward evidence with sensor data, and the procedures and protocols that might be in place to ensure the validity of citizen data. When citizens work with "indicative" air quality sensors that produce "just good enough data" (Gabrys, Pritchard and Barratt, 2016), however, we have found that the compliance-based approach of air quality monitoring offers just one particular way of investigating urban air pollution. Citizen air quality monitoring can demonstrate a much different set of attachments and concerns, as well as ways of working with data and evidence. Here, citizen data does not attempt to replicate or become an organ of expertise. However, it does differently constitute the problem of air pollution, which points to the plural urbanisms that converge through major and ongoing environmental crises such as air pollution. Data citizens, in this sense, are generated in relation to numerous forms of data and data practices, which become sites of collective making, interpretation and narration.

At the same time, the right to data and the right to the city become entangled with the right to produce evidence. Certain ways of establishing the facts of environmental problems are treated as more credible than others, with significant consequences for how cities develop and urban life is lived. Ruha Benjamin suggests that empiricism often only works for some, since no amount of evidence will be accepted if the "facts" challenge the status quo or are presented by marginal voices. As she writes, "The facts, alone, will not save us" (2016, 2). The data citizen, in this sense, is not automatically an enlightened political subject. Indeed, it could be an exclusive and exclusionary position, since data also requires environments of relevance in which to take hold. Whatever accomplishment citizen data makes in its observations, infrastructure and collective experiencing, in order for it to evidence environmental harm and realize improved environmental conditions it also needs to set in motion the worlds that enable that data to have effect. Effect in this sense is less about the success or failure of data, and more about the impasses that can arise when prevailing forms of political engagement break down or demonstrate their hollow promises. The practices of data citizens can in this way constitute processes of proposing and working toward these worlds where citizen data matter, and where the effects of data contribute to more livable processes of urbanization.

Conclusion: Propositions for citizen data and urban worlds

In a recent examination of sensing air quality in *Program Earth* (Gabrys 2016), I suggest that it might be possible to engage with data neither as free-floating facts, nor as the monolithic products of expertise, but rather as *creatures* that are constituted with and through environments of relevance. I draw on Alfred North Whitehead's discussion of creatures as the actual entities and occasions that concresce through processes and relations (Whitehead 1985). When we consider how monitoring practices are ways of creaturing air pollution data, then it is also possible to attend to the environments of relevance in which data is formed, the problems it responds to and is attached to, and its importance for those who generate this evidence.

Creaturing is a process whereby data can come to matter. But as I suggest here, the different creatures of air pollution data can also create sites of struggle. Which data matters, and which urban worlds are sustained? Which data is overlooked, and which urban worlds are extinguished?

While air pollution monitoring instruments can be made to align, more or less, to detect a similar pollutant level in space and time, the actual uptake, use, deployment of sensors as well as the generation of data veers into different directions when used by air quality officials for regulation, and when used by residents observing and documenting changes in the urban fabric. Not to attend to citizen data is to neglect urban dwellers' attachments to their cities, to the problems that matter in their urban lives, and to the practices whereby they document, analyze and communicate evidence that speaks to their concerns. To make expertise the only register for producing legitimate data is to forgo and forget the importance of the environments that sustain data and allows it to have effect. It is also to suggest that an annual average calculated to comply with a regulatory guideline is the only way to organize the problem of air pollution—as well as the only way of considering how to create possible preventative and mitigating actions. To adhere to one expert version of collecting data and making facts is also to miss the question of which problems these facts pertain to, and which worlds they sustain (see also Stengers 2011).

It is possible *both* for experts' data indicating that annual-mean levels of $PM_{2.5}$ are 19 $\mu g/m^3$ and for citizens' data indicating specific patterns of elevated emissions when viewed as 1-hour and 24-hour datasets to be "accurate." Each of these forms of data takes hold and gains relevance in distinct environments, and as specific responses to environmental problems. If a more pluralistic ontology of data were to be realized, then both—and more—of these creatures of data would need to be recognized as relevant to our inundated urban habitats. Indeed, the very qualities of expertise could begin to shift and respond along with the environmental conditions that are meant to be governed toward more collective projects, which might be better addressed through multiple urban experiences and data. Here is where data citizens materialize as figures constituted through the relations and communities in and on behalf of which evidence would be mobilized.

Such pluralistic ontologies extend from data and the problems it responds to, as well as the modes of citizenship and rights-in-the-making that might be materialized through data practices. No singular figure of the data citizen concretizes here. These are, as Berlant has suggested, proliferating forms of citizenship, since they are tied to the worlds that are endured, narrated, created and hoped for. Drawing on, yet also critically interrogating Lefebvre's well-known concept of the right to the city, while also gathering theoretical resources from Berlant's "cruel optimism," I have investigated citizen data as it is constructed and mobilized in and on behalf of urban worlds. Citizens who collect or analyze data might register new and significant observations, but these forms of evidence might not make a dent in political or regulatory processes. In this sense, rights to data are not easily configured through clear codes of access and use, since data might be "open" but only certain groups are able to mobilize or make claims with such data, often in relation to other data

sources and with access to particular trajectories of power. The right to the city, as expressed through citizen data collection, can be a project that is undertaken through struggle, and that falls flat if political environments and relations do not exist for building on that struggle.

As noted in the introduction to this chapter, an increasing amount of legislation is being enacted in order to protect citizens' rights in relation to data, whether through tracking, the right to be forgotten, the right to open data, the right to transparency or more. However, the generation of citizen data through citizen sensing technologies raises different sorts of issues that might on the one hand be more aligned in some ways with the expansive version of rights articulated by Lefebvre in the right to the city. On the other hand, Berlant's work is also instructive as it suggests that proliferating modes of citizenship are indications of different experiences that will inform how rights in the making are taken up, if at all, as well as the struggles they produce. These different affective engagements are productive of different ways of being in the world, as they make different worlds. The question of rights then traverses through to the question of worlds. With rights in the making, what sorts of worlds are also in the making? If rights are no longer an adequate description of the work that data citizens undertake, then how do these practices generate distinct modes of political engagement? Citizen data practices undertaken in relation to urban environmental problems raise this challenging set of questions. People take up devices and make their own data and analyze a range of datasets, motivated often by their concerns about unjust process of urbanization to which they have no official rights. Rights such as the right to clean air might exist in some cities and countries, but these rights are frequently not observed. Interventional citizen data practices potentially reinvent the terrain of rights—how they are formed, expressed, transformed, claimed or abandoned. Such data practices form along with political subjects and collectives that are in search of more livable urban worlds, but which rights do not fully support.

Acknowledgments

Thanks are due to Helen Pritchard, Lara Houston, Lau Thiam Kok, Benjamin Barratt, Khadija Jabeen, Sarah Garcin, Raphael Faeh, Francesca Perona and Adrian McEwen for contributing to the collaborative research, design, calibration and development of the citizen sensing technologies used in this research. Thanks are also due to the participants of the urban sensing project, and to the organizations that have hosted workshops and events, including Deptford Folk, Deptford Neighbourhood Action, Peyps Estate, Crossfields Estate, Voice for Deptford, APT Gallery, New Cross Gate Trust, Deptford Lounge Library and New Cross Learning.

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 313347, "Citizen Sensing and Environmental Practice: Assessing Participatory Engagements with Environments through Sensor Technologies."

Note

1 As reported in "India: Health Emergency Declared as Toxic Air Shrouds New Delhi," *Democracy Now* (8 November 2017), www.democracynow.org/2017/11/8/headlines/india_health_emergency_declared_as_toxic_air_shrouds_new_delhi. Based on this story, it is unclear which pollutants measured "999" on the Air Quality Index (AQI). The AQI is available at https://aqicn.org/city/delhi.

Bibliography

Arendt, Hannah. 1951. *The Origins of Totalitarianism*. New York: Harcourt Brace.

Balibar, Étienne, Barbara Cassin and Sandra Laugier. 2014. "Praxis." In *Dictionary of Untranslatables: A Philosophical Lexicon*, edited by Barbara Cassin, 820–832. Princeton, NJ: Princeton University Press.

Benjamin, Ruha. 2016. "Racial Fictions, Biological Facts: Expanding the Sociological Imagination through Speculative Methods." *Catalyst: Feminism, Theory, Technoscience* 2 (2): 1–28. doi:10.28968/cftt.v2i2.88.

Berlant, Lauren. 2007. "Citizenship." In *Keywords for American Cultural Studies*, edited by Bruce Burgett and Glenn Hendler, 37–42. New York: New York University Press.

Berlant, Lauren. 2011. *Cruel Optimism*. Durham, NC: Duke University Press, 2011.

Berlant, Lauren, interviewed by David K. Seitz. 2013. "On Citizenship and Optimism." *Society + Space*, March 22. http://societyandspace.org/2013/03/22/on-citizenship-and-optimism.

Berlant, Lauren. 2016. "The Commons: Infrastructures for Troubling Times." *Environment and Planning D: Society and Space* 34 (3): 393–419. doi: 10.1177/0263775816645989.

Citizen Sense. https://citizensense.net.

Citizen Sense. 2017. *Deptford Data Stories*, November 14. https://citizensense.net/data-stories-deptford.

Crawford, Margaret. 2011. "Rethinking 'Rights,' Rethinking 'Cities': A Response to David Harvey's 'The Right to the City.'" In *The Right to the City*, edited by Zanny Begg and Lee Stickells, 33–37. Sydney: Tin Sheds Gallery.

Das, Pamela and Richard Horton. 2017. "Pollution, Health, and the Planet: Time for Decisive Action." *The Lancet* 391 (10119): 407–408. doi: 10.1016/S0140-6736(17)32588-6.

Data for Black Lives. http://d4bl.org.

DeGooyer, Stephanie, Alastair Hunt, Lida Maxwell and Samuel Moyn. 2018. *The Right to Have Rights*. London: Verso.

Gabrys, Jennifer. 2016. *Program Earth: Environmental Sensing Technology and the Making of a Computational Planet*. Minneapolis, MN: University of Minnesota Press.

Gabrys, Jennifer, Helen Pritchard and Benjamin Barratt. 2016. "Just Good Enough Data: Figuring Data Citizenships through Air Pollution Sensing and Data Stories." *Big Data & Society* 3 (2): 1–14. doi: 10.1177/2053951716679677.

Gregory, Judith and Geoffrey Bowker, 2016. "The Data Citizen, the Quantified Self, and Personal Genomics." In *Quantified: Biosensing Technologies in Everyday Life*, edited by Dawn Nafus, 211–226. Cambridge, MA: MIT Press.

Harvey, David. 2008. "The Right to the City." *New Left Review* 53: 23–40.

Houston, Lara, Jennifer Gabrys and Helen Pritchard. 2019. "Breakdown in the Smart City: Exploring Workarounds with Urban Sensing Technologies." Citizen Sense Working Paper.

Isin, Engin and Evelyn Ruppert. 2015. *Being Digital Citizens*. London: Rowman & Littlefield.

Iveson, Kurt. 2013. "Cities within the City: Do-It-Yourself Urbanism and the Right to the City." *International Journal of Urban and Regional Research* 37 (3): 941–956. doi:10.1111/1468-2427.12053.

Kravets, David. 2017. "Law Making It Illegal to Collect Data, Photo of Open Land Hangs in Balance." *Ars Technica*, September 11. https://arstechnica.com/tech-policy/2017/09/ag-gag-law-gets-taken-to-the-slaughterhouse.

Lefebvre, Henri. 1996. *Writings on Cities: Henri Lefebvre*. Selected, translated and introduced by Eleonore Kofman and Elizabeth Lebas. Oxford: Blackwell Publishers.

Lewisham Council. www.lewisham.gov.uk.

London Datastore. https://data.london.gov.uk.

Mitchell, Don. 2003. *The Right to the City: Social Justice and the Fight for Public Space*. New York: Guildford Press.

Petrescu, Doina and Kim Trogal, eds. 2017. *The Social (Re)Production of Architecture: Politics, Values and Actions in Contemporary Practice*. Abingdon: Routledge.

Pidot, Justin. 2015. "Forbidden Data: Wyoming Just Criminalized Citizen Science." *Slate*, May 11. www.slate.com/articles/health_and_science/science/2015/05/wyoming_law_against_data_collection_protecting_ranchers_by_ignoring_the.html.

Purcell, Mark. 2002. "Excavating Lefebvre: The Right to the City and Its Urban Politics of the Inhabitant." *GeoJournal* 58 (2–3): 99–108.

Ruppert, Evelyn, Engin Isin and Didier Bigo. 2017. "Data Politics." *Big Data & Society* 4 (2): 1–7.

Shaw, Joe and Mark Graham. 2017. "An Informational Right to the City? Code, Content, Control, and the Urbanization of Information." *Antipode* 49 (4): 907–927.

Stengers, Isabelle. 2011. *Thinking with Whitehead: A Free and Wild Creation of Concepts*. Translated by Michael Chase. Cambridge: Harvard University Press.

United Nations. 2017. *New Urban Agenda: Habitat III*. Quito, Ecuador. http://habitat3.org/wp-content/uploads/NUA-English.pdf.

Whitehead, Alfred North. 1985. *Process and Reality*. New York: The Free Press.

Wynter, Sylvia and Katherine McKittrick. 2015. "Yours in the Intellectual Struggle: Sylvia Wynter and the Realization of the Living." In *Sylvia Wynter: On Being Human as Praxis*. Durham, NC: Duke University Press.

14

DATA RIGHTS

Claiming privacy rights through international institutions

Elspeth Guild

Introduction

This chapter examines the coming into existence of the data citizen. The data citizen is more than a data subject who is merely the object of state measures to protect the subject's right to privacy. The data citizen is an actor who is entitled by reason of national but most importantly international law to respect for his or her privacy. The data citizen is a rights holder in international law regarding his or her personal data. The data citizen is a citizen in so far as he or she is entitled to a bundle of rights in the Marshallian sense (Marshall and Bottomore 1992). The source of these rights may be national law, but that national law must be in conformity with international law, which creates the citizen's right to privacy. This has two consequences – firstly, the data citizen is not a citizen because a state has conferred on him or her that status. He or she is a citizen through the claim to rights in international law. Secondly, the data citizen is contesting and seeking to establish his or her citizenship right to privacy and is using the intersection of international and national law as a nexus through which to achieve these claims. Through these challenges, controversies and struggles, the data citizen is under construction. This process is likely to continue for at least the next three decades.

By focusing on the emergence of the data citizen, the rather sterile legal debate on the relationship between privacy and data protection falls away. Instead, as the data citizen comes into existence, he or she does so by reason of the struggle for control of his or her privacy against both public and private actors. In legal language, this struggle is often framed through the human rights principle of consent to use of personal data. A data citizen's personal data belongs to him or her for ever, no matter who has collected it. Every time a public or private enterprise wishes to use the data citizen's personal data, it must seek the consent of the data citizen. That consent must be expressed in full knowledge of the exact use that will be made of his or her

personal data. A data citizen cannot waive his or her right to consent regarding what is done with his or her data, or be coerced into doing so. A state can only interfere with the data citizen's personal data where it can justify that interference on the basis of the exceptions set out in international law (for instance in pursuit of crime, terrorism, etc.) and subject to all the limitations that apply to the exception such as justification, judicial oversight etc. This chapter examines how this data citizen is coming into existence notwithstanding the profound obstacles that both some private sector actors and some states are putting in the citizen's way.

Why data citizens?

Revelations about the extent of state surveillance of electronic communications in June 2013 by Edward Snowden had a substantial impact on civil society and state authorities in Europe. The view that mass surveillance of electronic communications is an interference with a profound element of privacy of the person was widely expressed. Privacy and its regulatory counterpart, data protection, have traditionally been defined in the context of citizen's rights within liberal democracies (Bennett and Raab, 2006). The definition of privacy and the categorisation of personal (and sensitive personal) data for the purposes of state protection, while varying according to national sensitivities, has nonetheless been focused on and expressed in the language of the state-citizen relationship. The formulation of privacy and data protection as components of international human rights has taken place within a framework that does not differentiate on the basis of nationality (Chander and Lê 2014, 677–739). International obligations (such as the International Covenant on Civil and Political Rights) and regional obligations (such as the European Convention on Human Rights and the EU Charter of Fundamental Rights) are all coached in the language of human rights, which is to say, applicable to everyone irrespective of citizenship. Yet, if one looks at the controversies about protection of privacy and personal data obligations, almost all of them revolve around claims by citizens against their own state. While a citizen-foreigner divide is not part of the human rights commitments of states in respect of privacy (or data protection) in practice, it is almost exclusively citizens who challenge state practices in respect of this human right. This controversy is revealed by the language used: (a) data "subjects" – the humans whose entitlement towards their personal data and privacy is passive, regulated by the state, and only transformed into an active right by constitutional rights for citizens (b) data "rights" – the struggle of people to have definitive rights over their personal data and thereby become actors (the antithesis of the argument made in US law that personal data belongs to the entity that collected it) irrespective of what constitution applies and (c) data citizens – an incipient category founded in international human rights law that entitled people to human rights over their personal data, irrespective of their state citizenship, and which is characterised by a genuine right of the citizen to consent or refuse to consent to the use of his or her personal data.

On entering this territory of controversy, it is important to remember that the right to privacy of everyone (the language of international human rights) is

a human right. It takes primacy over all claims, public or private sector, to use personal data. These actors can only use personal data of the data citizen on two conditions – either the data citizen has agreed and given valid consent, or there is an argument in public policy (for instance the investigation of a crime), which is sufficiently clear and precisely articulated to permit an exception to be made and the consent of the data citizen to be dispensed with. The duty of states to guarantee to data citizens the protection of their personal data is the way in which states are required to deliver the data citizen's privacy. Data protection is a necessary tool in the delivery of the privacy right. It is not some separate category only tangentially related to the right to privacy. It is the obligation on states to deliver on the privacy right of the data citizen.

Data citizen

The transformation of the capacity to capture and use electronic communications is creating a data citizen whose rights and obligations do not derive exclusively from the state. The reason for this is the profoundly transnational nature of the transmission of personal data across international borders. The citizen of one country who sends an email to a colleague in the next office cannot rely on national law on privacy and its implementation through data protection by his or her own state. This is because that email may well cross many international borders as it is chopped into pieces, transmitted around the world by whatever is the fastest route and reassembled, and delivered to the colleague in the next office. The national law of one country cannot necessarily protect the privacy of citizens of that state as the internet does not respect national borders in the transmission of personal data of citizens. This process means that personal data may travel through the jurisdictions of many different countries before arriving at its destination. Each country will have different rules on privacy and data protection. A number of effects that are commonly bundled up in the concept of globalization have muddied the waters. First, within the private sector, personal data of individuals is often categorised as consumer data rather than citizens' data. The private sector, particularly in the form of transnational companies, frequently want to treat this personal data in countries other than those where the data was collected. This can be because the data citizen's rights are better protected in one country than in another and transnational companies may seek to avoid strict rules on consent by using jurisdictions of weak or non-existent data protection to manipulate their personal data information. Thus, while the individual may remain a citizen in his or her own country, his or her personal data that has been collected by a private sector body may be sent for processing to some other country. While for the purposes of the company, the individual remains a consumer, for the purposes of the individual, he or she has moved from being a citizen in his or her own state vis-à-vis the personal data that was collected to being a foreigner vis-à-vis the country where his or her personal data is being processed. Differences in privacy and data protection laws at the national level and in the extent to which states have adhered to international

human rights obligations relating to privacy and personal data become central for the individual as regards what happens to his or her privacy (Bennett 2017).

Secondly, people move. According to the UN World Tourism Organization (UNWTO 2018), there were more than 1 billion tourists who moved from one country to another in 2016 (counted on the basis of arrivals). The vast majority of these people are citizens of their country of departure, but on leaving become foreigners in the country where they arrive. Their travel activities transform them from citizens into foreigners and back again every time they move from home abroad and back. People who travel across international borders have become a category of persons about whom very substantial amounts of information are gathered. The private sector collects information from such persons for its own marketing and monitoring purposes; these are customers. The sector also collects information for states under a variety of obligations, though generally this is through personal data sharing arrangements about data that the sector already collects.

The best known is the Passenger Name Record (PNR) data sharing duty, which a small but growing number of governments (so far) impose on all airlines carrying people to their country or travelling through their airspace.[1] The EU is establishing a similar programme (Mitsilegas, Valsamis, and Vavoula 2017, 232). Here, the personal data is about individuals. They have to provide it to airlines in the context of booking flights, receiving boarding cards, etc., though most of those individuals were citizens in their own state and maybe contracting with local companies. The data is made available to state authorities not only elsewhere in the world but also in their own state; states claim the right to treat that personal data as belonging to foreigners and use it for intrusive automated processing purposes (though there will also be own nationals among them).

Thirdly, international travel results in the collection and use of personal data in other ambiguous citizen-foreigner contexts. For instance, anyone who is subject to a mandatory visa requirement to travel to a specific country will be obliged to provide information to the destination country in order to obtain a visa.[2] The information that individuals may be required to provide about themselves in order to obtain a visa can be quite extensive and intrusive. For example, for those who must obtain a Schengen visa to visit any of the EU states that participate in the Schengen area, consulates may require that they produce wages slips for a period of months before the intended departure, information about family members and their whereabouts, biometric data including fingerprints etc. All this personal information is then stored in the Visa Information System and made available to law enforcement authorities in the EU (Schengen) Member States. While the personal data will be collected from the individual while he or she is a citizen in his or her own country (and thus at least potentially subject also to national laws), it is collected by the authorities of a foreign state, transmitted to a database in the EU and made available to many law enforcement authorities of countries with which the individual may never have had any contact at all. The act of seeking a visa for a short stay in a foreign country is the trigger that the foreign country uses to collect, process and share that personal data with a wide range of actors and other states.

Fourth, when people move for longer periods, for instance to study in a foreign country or work there for a while, they will need to provide much more detailed personal data to the immigration authorities of the host state. Immigration forms for residence on these grounds often run to many pages and must be supported by sensitive personal data about finances, health status etc. However, that personal data may or may not be secure. For example, the UK forms require the individual to consent to the use of his or her data in the following terms:

> I understand that all information provided by me to the Home Office will be treated in confidence but that it may be disclosed to other government departments, agencies, local authorities, the police, foreign governments and other bodies for immigration purposes or to enable them to perform their functions, and that, if such bodies provide the Home Office with any information about me which may be relevant for immigration purposes, it may be used in reaching a decision on my application.

Under a UK-US agreement, the US authorities are entitled to access to all the information that the individual provided to the UK authorities for the consideration of his or her immigration application. Data citizens: citizens of what state?

How do states and data citizens interact on the basis of their relationship – state/citizen, foreigner/state? How do people claim data rights that transcend the traditional citizen foreigner divide, and where do they look for these new rights to transform themselves into data citizens? I will address this question from two perspectives. First, how have international human rights obligations with respect to privacy or personal data protection become the place to search for data rights outside the framework of state constitutions for their citizens? And second, what are the state practices at play that seek to blur the lines and undermine their own state constitutional guarantees to citizens? I will examine these questions under the following categories:

1. Privacy and personal data protection and the International Covenant on Civil and Political Rights (ICCPR) – is everyone the data citizen?
2. Personal data sharing agreements in the area of borders and immigration among states versus the data citizen.

In examining these two categories, I will return to some of the examples I have referred to in this introduction and provide a fuller and more in-depth examination of the issues at stake.

Data citizens and international human rights: the ICCPR

The Snowden revelations regarding mass surveillance have not only had very substantial political repercussions over 2013 and into 2014, but have also raised profound legal questions. Among these is the question of the protection of the

privacy and data of citizens in comparison with the protection of the same rights in respect of foreigners. One of the most striking aspects of the response of the US Government to the concerns articulated by governments of other countries about the revelations of US mass surveillance internationally has been the US President's assurance to US citizens, but only to US citizens, that their privacy will be fully protected. While the privacy of foreigners will be augmented, the (then) President made it very clear that foreigners cannot rely on the same privacy protections as US citizens. The reason for this is because in the logic of the President's speech, foreigners are more of a security risk to the USA than its own citizens.

Specifically, the then US President Barack Obama's speech of 17 January 2014 stated that "the legal safeguards that restrict surveillance against U.S. persons without a warrant do not apply to foreign persons overseas" (White House Press Office 2014). Here, the citizen-national constitutional rights framework of privacy is clearly demonstrated. The US Government's reliance on law as national constitutional law is clearly drawn. In pursuit of these constitutional issues related to the Snowden revelations, the US President clarified how the state would protect the constitutional rights of US citizen (at least those who stay in the USA). The steps that the President promised in the speech included

> we will reform programs and procedures in place to provide greater transparency to our surveillance activities, and fortify the safeguards that protect the privacy of U.S. persons . . . we will provide additional protections for activities conducted under Section 702, which allows the government to intercept the communications of foreign targets overseas who have information that's important for our national security. Specifically, I am asking the Attorney General and DNI [Office of the Director of National Intelligence] to institute reforms that place additional restrictions on government's ability to retain, search, and use in criminal cases communications between Americans and foreign citizens incidentally collected under Section 702."

The fracture of the national constitutional logic on the right to privacy appeared at a slightly later stage in the President's speech when he stated, "I have taken the unprecedented step of extending certain protections that we have for the American people to people overseas. I've directed the DNI, in consultation with the Attorney General, to develop these safeguards, which will limit the duration that we can hold personal information, while also restricting the use of this information" (Ibid). This nod to concerns of governments elsewhere in the world to the surveillance of their citizens is couched in the terms of noblesse oblige; the unprecedented quality of the US Government's commitment to protect the privacy of foreigners is not acknowledge by the US authorities as founded in any legal obligation to do so. It is unprecedented because it goes beyond the legal obligations of the US government under its national constitution. The qualification of the protection of the privacy

of foreigners as "unprecedented" reveals a profound refusal to accept the premise that the international human rights obligations of states (including the USA) affect the way in which states act towards personal data of foreigners.

Yet, three interconnected but separate human rights issues arise as regards mass surveillance. The first, which is the most fundamental but is the most frequently ignored, is the prohibition on arbitrary or unlawful interference with privacy,[3] alternatively formulated as the right of every person to respect for his or her private and family life.[4] The second, which is generally the subject of more substantial political and media noise, is the duty of states to protect personal data. The noise around personal data protection general depends on the differing national legislation that is designed to give effect to the right to privacy. The third is the right of expression, the so-called "chilling factor" that knowledge of mass surveillance has on people's willingness and ability to express their opinions freely (La Rue 2011). This human rights violation related to mass surveillance has been most forensically analysed in the contest of academic work on effect of surveillance by the former East German Stasi and Romanian Securitate on their people's right of expression.[5] The conclusion of these works is that knowledge of or the well-founded suspicion that state authorities are listening into private conversations and communication has profound consequences on the ability of people to express themselves freely.

Those political actors who have an interest in promoting the legality of mass surveillance usually put forward two arguments. The first is that national and international security is always an exception to both the duty of every state to respect people's privacy and the duty to protect personal data. Where national and international security interests are at stake, states are not only entitled to rely on the qualification of the rights but also may have differing standards, which are nonetheless consistent with the margin of appreciation applicable to the rights.[6] This is the most trenchantly defended of arguments, as when this one falls away, those actors seeking to justify mass surveillance find themselves on weaker legal ground (Harris 2010). The second argument is that states' obligations to protect personal data are subject to very different rules and requirements according to the political preferences of different states. Thus, as there is no harmonization of the specific rules as to what is acceptable data protection internationally, states that are exercising their national and international security prerogatives only need to fulfil their own national data protection rules.[7] For the moment, the response to the chilling effect on the right to freedom of expression has been rather muted at least by those political actors interested in promoting mass surveillance as a necessary security tool.Yet, the right to respect for a person's privacy is an overarching international human right. It is found in the UN's Universal Declaration of Human Rights 1948[8] and its constraining legal form is contained in the UN's International Covenant on Civil and Political Rights 1966.[9] Any interference with the privacy of a person must first and foremost be subject to the consent of that person. That consent must be informed, that is to say the individual must know exactly what he or she is consenting to, how his or her personal data may be used and the precise limits to that use. Uninformed consent is not consent. The right to consent

or refuse use of personal data belongs to the individual not the state. Thus, any interference without consent is an exception and as such must be strictly limited. Where the state seeks to interfere with that right and to collect and use personal data, which constitutes an intrusion into the privacy of the person concerned, such an interference must be justified by the state authorities. First it must be permitted by law and that law must be sufficiently clear and public that everyone can know what the law states and how to adjust their behaviour accordingly. Any exception permitted by law to a human right must be interpreted narrowly. It must have a legitimate objective and be necessary to achieve that objective only. There must be no alternative, which would be less intrusive into the life of the person that could instead be used. There must be judicial oversight of any state interference and a person affected by an interference must have access to justice to challenge that interference (Bennett 2011). Mass surveillance by its very nature is not targeted at any person specifically thus the possibility to justify the interference with the privacy of any person individually is an exceedingly difficult task. Where such mass, weakly targeted surveillance techniques have been used in Europe, the European Court of Human Rights has found them inconsistent with the right to respect for privacy. The core problem is that mass surveillance is, by definition, arbitrary.[10] The challenge for the data citizen is twofold: first, to establish his or her existence as a citizen with rights in the transnational world of personal data transmission, and second, to push the international community to take responsible for the protection of the data citizen's data through concrete measures clarifying the data citizen's right to his or her privacy and achieving agreement and practical guidelines and roadmaps for states to deliver the data rights of this new kind of citizen.

Moving then from the state of human rights to the political struggle regarding mass surveillance, clearly the US authorities are faced with a dilemma in international human rights law, an area about which they have always been rather wary. The 1950s approach to international human rights law was to claim that the instruments do no more than set out principles and are not 'real' law in any significant way and are certainly not available for people to rely upon (Ishay 2008). This political position has been undermined by the development of very precise international obligations, the establishment of Treaty Bodies with jurisdiction to receive and adjudicate on complaints by individuals regarding alleged breaches of their international human rights and the embrace of international human rights law by national courts (Falk 2002). The principles approach to international human rights law is no longer tenable, it is a fig leaf deployed occasionally by states seeking to act arbitrarily (Ghandhi 2012). As the Snowden revelations became increasingly politically salient among international issues in 2013–4, a number of states, primarily led by the Brazilian and German authorities began to address the issue of how to deal with US mass surveillance and interception of communications. There was much discussion about bilateral negotiations and unilateral action, for instance building new cables that avoid US territory (Blau 2014, 14–16). However, it was rapidly evident that bilateral and unilateral approaches were not going to be satisfactory. In Europe, the UK authorities were carrying out mass surveillance for their

US counterparts and others – the so-called Five Eyes, the surveillance alliance of Australia, Canada, New Zealand, the USA and the UK (Greenwald, Poitras and MacAskill 2013) – yet were not only members of the Council of Europe but also of the European Union. This was only one example of the problem of unilateral or bilateral approaches. Clearly only multilateral efforts were likely to bring results where the weight of the USA and some of its collaborators could be counterbalanced by a loose alliance of other states. As soon as the issue is defined in this way, the obvious venue to commence a response is the UN General Assembly and the territory on which to prepare the response is international human rights obligations – the prohibition of arbitrary interference with peoples' privacy.

This is the road that the Brazilian and German authorities followed. By August 2013 the moves were afoot for a resolution of the General Assembly. Five non-governmental organizations were closely linked with the efforts; Access, Amnesty International, Electronic Frontier Foundation, Human Rights Watch and Privacy International were also applying pressure for a strongly worded resolution. The Brazilian and German authorities were by no means alone in their efforts to achieve agreement of a UN General Assembly Resolution. Many smaller states, most notably Austria, Hungary, Liechtenstein, Norway and Switzerland, but also others, very strongly supported the work from the beginning even seconding staff to assist with the workload. The matter was assigned to the General Assembly's Third Committee and it is there that the tense negotiations on the wording of the Resolution took place. A text was adopted on 26 November in the Third Committee and on 18 December 2013 it was adopted without vote in the General Assembly of the UN (Carrera, Guild and Parkin 2014).

The Resolution is based on the right to respect for privacy in the Universal Declaration and the ICCPR with specific reference to the prohibition on arbitrary interference. It ties the right to privacy to the right to freedom of expression – if people are subject to mass surveillance, they are no longer able to express themselves freely. The preamble to the Resolution insists on the negative impact that surveillance and interception of communications, including extraterritorial surveillance and interception, on a mass scale has on the exercise and enjoyment of human rights. The Resolution calls upon states to respect the right to privacy and prevent violations; to review their procedures, practices and legislation regarding surveillance of communications, their interception and collection of personal data, including mass surveillance, interception and collection with a view to upholding the right to privacy and ensuring the full and effective implementation of all their obligations under international human rights law and to establish or maintain independent, effective domestic oversight mechanisms capable of ensuring transparency and accountability of state's actions.

The UN Human Rights Council (composed of 47 states elected by the General Assembly) has also already engaged with the issue. The High Commissioner noted at that meeting that the threat that mass surveillance poses to human rights is among the most pressing global human rights situations today. Many state representatives present at that session had regard to the report of UN Special Rapporteur on the

promotion and protection of the right to freedom of opinion and expression. La Rue (2011) had already outlined many dangers of state surveillance and its impact of free speech. What is perhaps surprising is that the September 2013 meeting of the Human Rights Council received so little press coverage. The meeting was well attended by state representatives. The discussions were incendiary as many state representatives attended the meeting with statements of condemnation of mass surveillance and interception of communications already prepared and agreed with neighbouring states on whose behalf they were mandated to speak. While one might well expect the German representative to present a text on behalf of Austria, Hungary, Liechtenstein, Norway, and Switzerland it is perhaps less obvious that Pakistan, speaking on behalf of Cuba, Venezuela, Zimbabwe, Uganda, Ecuador, Russia, Indonesia, Bolivia, Iran and China would also present an agreed text condemning the practices. While the counter move particularly in respect of this second set of countries is usually to attack them on the basis of their internal practices of surveillance and suggest, if not accuse them of, hypocrisy, the fact of the intervention nonetheless must be noted and the possibility that a group of states with serious disagreements among themselves would choose common ground on this subject.

What the debate in 2013 and 2014 reveals is that the division of privacy rights along the lines of citizens/foreigners is increasingly challenged on the basis of universality of human rights. The categorisation of privacy as an issue exclusively determined on the basis of constitutional relationships of citizens with their states has now been fundamentally challenged by data rights as human rights. The UN General Assembly has rejected the citizen-foreigner distinction, arguing that privacy is an international human right guaranteed by international instruments that must be respected by all signatory states. The attempt by the US authorities to frame the protection of the privacy of foreigners as an act of generosity by the US authorities unrelated to US state human rights obligations has been rejected by the UN General Assembly. Further, those states that carried the debate in the General Assembly were in no doubt that the right to privacy of all people, whether citizens or foreigners, vis-a-vis the state collecting the private data has extraterritorial effect. Where a state exercises its jurisdiction by collecting personal data anywhere in the world, that exercise of jurisdiction, implied or explicit, carries with it the human right prohibition on arbitrary interference with privacy.

The immediate outcome has been the creation of a new post at the UN level of Special Rapporteur on Privacy, and the appointment of Professor Joseph Cannataci of Malta in July 2015. The mandate (Resolution 28/16) given to the Special Rapporteur by the Human Rights Council is:

(a) To gather relevant information, including on international and national frameworks, national practices and experience, to study trends, developments and challenges in relation to the right to privacy and to make recommendations to ensure its promotion and protection, including in connection with the challenges arising from new technologies;

(b) To seek, receive and respond to information, while avoiding duplication, from States, the United Nations and its agencies, programmes and funds, regional human rights mechanisms, national human rights institutions, civil society organizations, the private sector, including business enterprises, and any other relevant stakeholders or parties;
(c) To identify possible obstacles to the promotion and protection of the right to privacy, to identify, exchange and promote principles and best practices at the national, regional and international levels, and to submit proposals and recommendations to the Human Rights Council in that regard, including with a view to particular challenges arising in the digital age;
(d) To participate in and contribute to relevant international conferences and events with the aim of promoting a systematic and coherent approach on issues pertaining to the mandate;
(e) To raise awareness concerning the importance of promoting and protecting the right to privacy, including with a view to particular challenges arising in the digital age, as well as concerning the importance of providing individuals whose right to privacy has been violated with access to effective remedy, consistent with international human rights obligations;
(f) To integrate a gender perspective throughout the work of the mandate;
(g) To report on alleged violations, wherever they may occur, of the right to privacy, as set out in article 12 of the Universal Declaration of Human Rights and article 17 of the International Covenant on Civil and Political Rights, including in connection with the challenges arising from new technologies, and to draw the attention of the Council and the United Nations High Commissioner for Human Rights to situations of particularly serious concern;
(h) To submit an annual report to the Human Rights Council and to the General Assembly, starting at the thirty-first session and the seventy-first session respectively.

The Special Rapporteur presented his report (A/HRC/34/60) to the Human Rights Council in February 2017 calling for privacy-friendly oversight of government surveillance.

Personal data sharing among states – hastening the emergence of the data citizen?

As people move between countries, they find themselves straddling between the status of citizen and foreigner. This is a moment of particular vulnerability when people find themselves subject to a variety of state acts that would not be legal within a state but are countenanced at the border. The detention of people at the border is one of the areas that has received substantial attention over the past decade. While citizens cannot be prevented from entering their state of nationality, foreigners (who only a few hours earlier may have been citizens in their own state) have not such guarantee. Thus, to prevent them from making an unauthorised

entry onto the territory of a state, wide detention powers have been adopted by many liberal democracies (Cornelisse 2011, 207–226; Wilsher 2012). In the context of border and immigration controls, because states have claimed the sovereign power to control their external borders, foreigners seeking to cross those borders may be required to provide very substantial amounts of personal data in order to justify to officials the reason for their admission. This power to require people to hand over large amounts of personal data in order to enter a state has been widely used liberal democracies including against the citizen of one another. In this context, states increasingly claim the right to require the individual to consent to the use of his or her personal data in ways that the state itself will determine after the fact. The idea that there is some purpose limitation regarding the use of personal data of foreigners who have been required to provide the data in the context of immigration procedures has not yet been accepted by many state authorities. As mentioned in the introduction, the UK authorities require anyone seeking a visa for short or long stay in the UK to "consent" to the UK authorities sharing their personal data not only with an unspecified wide range of actors within the state but also with the authorities of third countries. The wording of the UK consent does not even limit the exchange and use of personal data to immigration related purpose, as it includes providing personal data to foreign authorities to enable them to perform their functions, whatever those might be. This is an exceptionally widely formulated "consent". Indeed, it is so wide that it is hard even to categorise it as consent as the individual has virtually no idea nor any way of finding out what may become of his or her personal data, or with whom it might be shared.

Just when the world's data-related attention was fixed on the Snowden revelations in 2013, the UK and US authorities entered into an agreement to exchange personal data of people provided o them in the context of immigration procedures. A year earlier, the US authorities had entered into a similar agreement with their Canadian counterparts,[11] but for my purposes I will focus on the UK/US agreement not least because of the European considerations that arise.

The US/UK agreement is important as it permits the exchange of substantial amounts of personal data collected in the context of visa and immigration procedures between the two countries. As everyone who works in the immigration field is well aware, people must divulge enormous amounts of sensitive personal data in immigration procedures. This agreement establishes a framework for a new level of exchange of personal data between states, which is purported to be lawful.

The agreement was signed on 18 April 2013 in New Zealand.[12] There are four aspects of the Agreement that bear attention: (1) the definition of Information; (2) the scope and purpose of the agreement; (3) disclosure and use of data; (4) protections for the data subject.

Information

The definition of Information is a key part of the agreement as everything else depends on it. Information means data that a person provides to either of the

authorities (US or UK) for the purposes of: (1) authorisation for transit; (2) travel (e.g. visas, ESTAs etc.); (3) work (all categories); (4) residence (all categories); (5) citizenship applications (all types).

This includes personal data on admissibility, immigration or nationality compliance actions and or decisions. All this personal data is Information that may be shared between the parties. This covers just about everything that a person provides to the immigration authorities on either side of the Atlantic. For instance, in the context of family reunification applications, people frequently need to provide photographs of their marriages and lists of those who were present in order to satisfy the authorities that the marriage is genuine. This personal data, including personal data of third parties, would constitute Information for the purposes of the Agreement.

Scope and purpose

Personal data categorised as Information can be shared between the UK and US authorities for the following purposes: to enforce or administer immigration and nationality laws of either party; to facilitate decision-making on applications for transit, visas, admission, extension of stay, other immigration benefit, nationality or removal; to prevent, investigate or punish acts that would constitute a crime that would make the individual inadmissible or removable under the laws of either party. As commission of even fairly minor crimes can be a mandatory refusal ground for entry onto the territory of the UK (and the US) this is a rather wide catch-all provision.

This personal data can be provided either on a systemic search or a case-by-case one. So, the US authorities could ask for all personal data on, for instance, anyone who has received a specific kind of permission to remain in the UK leave and vice-versa in a systemic search (or all asylum seekers), or could ask for all personal data on a named individual.

The scope of the agreement is limited to personal data on non-British or US citizens, though the temporal element in respect of information on citizenship applications is entirely unclear. The UK authorities have also signalled that "The UK will hold limited, if any, Information about European Economic Area nationals and their family members due to their free movement rights under European Union (EU) law." This is rather weak limitation as there is no UK commitment *not* to share Information on EU citizens and their family members with their US counterparts. It is only a warning to the US authorities that the UK may not have very much on these people. Also, there is no limitation on the UK authorities seeking information on EU citizens and their family members from the US authorities.

Disclosure and use of data

The UK and US authorities agree to provide one another with Information so long as it is for one of the purposes set out earlier. They may disclose all the personal data to any of their domestic authorities, which can make out an argument that they

have a role in carrying out one of the purposes. This could include criminal justice authorities, intelligence and police. The parties agree not to disclose this personal data to any private party, the public, a foreign government, international organization or court without express consent of the other party. However, if the other party consents then they can share further the data with courts and other governments. The only limitation is that when the authorities are sharing personal data, they make their best efforts to ensure that the data is not disclosed to home authorities of refugees or persons with protection under the UN Convention against Torture (that is to say, disclose to the persecutors). If the person has not yet been granted refugee status or CAT protection then the parties should not share his or her personal data with the home state if it is "reasonably foreseeable" that the person will be granted international protection. The same goes for the asylum seeker's family members.

What is particularly interesting here is that it is clearly foreseen that personal data may be shared with the country of origin of the individual. While British and US citizens are excluded from the scope of the Agreement, data for instance on Canadian citizens can be shared between the two parties and with consent, to the Canadian authorities as well. A similar provision in the US-Canada Agreement would allow for instance for information on British citizens who have provided information to the Canadian authorities to be passed to the US authorities. With the consent of the Canadian authorities, the US authorities could then share that personal data with the UK authorities so long as there was no question of the individual being a refugee from the UK. In this way, the restriction on the sharing of personal data to that of foreigners can be undermined. Provided that states have agreements with a sufficient number of countries then the treatment of personal data of foreigners according to one set of rules that is particularly lax and become the norm even for the collection of personal data about the state's own citizens.

Thus, the principle implicit in the Agreement that personal data of citizens of the state parties must be subject to different rules than that of foreigners may become meaningless.

Where is the data controller to protect the rights of the data citizen?

There is not much regarding the data subject him or herself. The UK and US authorities confirm to one another in the agreement that they have systems whereby people can request access to their personal data and its correction or notation. They further assure one another that where their authorities refuse to give access to data or to correct it, the data subject can seek redress. The UK and US authorities permit themselves to retain all the personal data they have exchanged for as long as they think it necessary. They are only obliged to destroy personal data if it is not relevant to a purpose or erroneously provided. The wording of the Agreement makes it clear that no rights are conferred on the data subject as a result of it. Enforcement, for instance of the rights of the data subject is only available to one of the parties to the Agreement, not to those affected.

This type of agreement evidences the fragility of people when they are within the power of immigration and border authorities and are not citizens of the state. It also highlights the assumption that the privacy of citizens of states parties to such agreements should not be covered by its provisions because their privacy is entitled to a higher standard of protection than that of the foreigner. What happens in this context is that the personal data and privacy of the individual who has become a foreigner by travelling is available for use without the same strict protections, which would be applicable if the individual was a citizen in his or her own state. In the context of these Agreements, the only international commitment that is directly referred to is the Refugee Convention.[13] While that convention is incorporated into the Agreement, other human rights conventions are ignored.

However, it became clear to both parties to the agreement that the data citizen escapes the categorisation of foreigner or citizen because the data citizen is an incipient legal subject whose authority and legitimacy derives from international rather than national law. This change of perspective was incorporated into an exchange of notes to the agreement which by implementing arrangements permits the extension of its scope to nationals of the parties.[14] So, the data citizen comes into being. He or she is the citizen of a state, but for the purposes of the collection, storage, processing and sharing of his or her personal data, he or she is now a data citizen and will have to rely on supranational sources of law to protection his or her data rights.

Conclusion

The transformation of the way personal data moves around the world is heralding the emergence of a data citizen – everyone, in the language of international human rights law – who is entitled to privacy and as a consequence the right to consent to (or refuse) the use of his or her personal data by pubic and private actors. It is evident that a number of liberal democratic states consider that the privacy and personal data of foreigners does not deserve the same degree of protection as that of citizens. This is evident in the aftermath of the 2013/2014 Snowden revelations not only from the practices revealed but more importantly from the reaction by the US authorities. The issue of mass surveillance of personal data was framed as one in which the citizen foreigner distinction was of substantial importance. While the US authorities accepted their duty to respect the privacy of their own citizens and the legitimacy of concerns by their own citizens regarding surveillance practices on themselves, they refused to acknowledge the same right to foreigners. To the extent that foreigners enjoy privacy or protection of their personal data, this is by reason of the generosity of the US authorities.

This approach to the citizen/foreigner divide was challenged by a number of countries that sought to express the right to privacy as an essential element of the prohibition on arbitrary interference with privacy contained in the ICCPR. Thus, the venue for the struggle was UN human rights obligations and their capacity to efface the difference regarding privacy and protection of personal data between foreigners and citizens.

The second case study that I have examined to throw light on how the citizen/foreigner differentiation works in the area of privacy, is the treatment of personal data provided to national authorities in the context of immigration and border procedures. Here I analysed an agreement between the USA and the UK regarding the sharing of such personal data in order to understand how this personal data is perceived and treated. It is evident that the countries entering into the agreement considered that the privacy of their citizens was such that even where their citizens are foreigners within the territory of the other contracting party, their privacy should be respected and their data not made subject to the data sharing rules. On the other hand, virtually all personal data provided by foreigners to the authorities of the parties is eligible for sharing. That sharing can be with a rather wide range of state actors and can extend to third countries on the basis of state consent. It is evident from the agreements that the states entering into them do not consider themselves constrained by the same rules about privacy and personal data protection that apply to citizens when they are dealing with the data of foreigners. But in recognition of the difficulty of separating the citizen and the foreigner in the data world, the agreement was extended to include all persons – the new data citizen as the object of state surveillance.

The data citizen will need to emerge from international human rights obligations of states. This requires the active participation of international organisations and civil society bodies in promoting the existence of this new citizen. Effective rights that characterise citizenship will only emerge in this field with the effective contestation of the claimed legitimacy of certain states to a monopoly over their entitlement to use everyone's personal data as they wish.

Notes

1 Australia, Canada, Japan, the European Union, the USA.
2 There are exceptions of course such as the Turkish visa system where foreigners buy the visa at the border when they arrive and thus provide very little information about themselves.
3 Article 17(1) International Covenant on Civil and Political Rights 1966 "No one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence, nor to unlawful attacks on his honour and reputation."
4 Article 8 European Convention on Human Rights.
5 cf. *The Lives of Others* (Sony Pictures Home Entertainment 2007); Deletant 1995; Dennis and Laporte 2003.
6 This doctrine has been much developed by the organs of the Council of Europe in the context of the ECHR (cf. Council of Europe 2014).
7 For an excellent analysis see Hosein and Palow 2013. For some of the counter arguments see Wright et al 2009.
8 Article 12: No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.
9 Article 17(1) ICCPR.
10 As in *Klass & ors v Germany*, 6 September 1978; *S & Marper v UK*, 4 December 2008.
11 Agreement between the Government of Canada and the Government of the United States of America for the Sharing of Visa and Immigration Information 13 December 2012; Canadian Treaty E105246.

12 Why there, one might wonder?
13 UN Convention relating to the status of refugees 1951 and its 1967 Protocol.
14 Exchange of Notes to amend the Agreement between the Government of the United Kingdom of Great Britain and Northern Ireland and the Government of the United States of America for the Sharing of Visa, Immigration, and Nationality Information, done at Queensland on 18 April 2013 signed 29-Sep-2016; published: Treaty Series 035/2016: CM9279 (554KB).

Bibliography

Anupam C. and U. P. Le. 2015. "Data Nationalism", *Emory Law Journal* 64(3), 677–739.

Bennett, C. 2011 "In defense of privacy: The Concept and the Regime". *Surveillance & Society*, 8(4), 485–496.

Bennett, C. 2017. "Redress, the International Protection of Privacy and National Security and Intelligence Agencies: The Role for an Ombudsperson". Available at SSRN: https://ssrn.com/abstract=3023624 or http://dx.doi.org/10.2139/ssrn.3023624

Bennett, C. and C. Raab. 2006. *The Governance of Privacy: Policy Instruments in Global Perspective*. Boston, MA: MIT Press.

Blau, J. 2014. "NSA surveillance sparks talk of national internets". *Spectrum, IEEE* 51(2): 14–16.

Carrera, S., E. Guild, and J. Parkin. 2014. "Who will monitor the spies?" *CEPS Commentary*. 8 January.

Chander, A. and Lê, U. P. (2014). "Data nationalism". *Emory Law Journal*, 64: 677.

Cornelisse, G. 2011. "Detention of foreigners" in *The First Decade of EU Migration and Asylum Law*. 207–225. Leiden: Martinus Nijhoff.

Council of Europe. 2014. "The Margin of Appreciation". Accessed at www.coe.int/t/dghl/cooperation/lisbonnetwork/themis/echr/paper2_en.asp on 20 February 2014.

Deletant, D. 1995. *Ceauşescu and the Securitate: Coercion and Dissent in Romania, 1965–1989*. Armonk, NY: M.E. Sharpe.

Dennis, M. and N. Laporte. 2003. *The Stasi: Myth and Reality; Themes in Modern German History*. London: Longman/Pearson.

Falk, R. A. 2002. *Human Rights Horizons: The Pursuit of Justice in a Globalizing World*. London: Routledge.

Ghandhi, S., ed. 2012. *Blackstone's International Human Rights Documents*. Oxford: Oxford University Press.

Greenwald, G., L. Poitras, and E. MacAskill. 2013. "NSA shares raw intelligence including Americans' data with Israel." Accessed at www.theguardian.com/world/2013/sep/11/nsa-americans-personal-data-israel-documents on 12 September 2013.

Harris, S. 2010. *The Watchers: The Rise of America's Surveillance State*. Penguin.

Hosein, G. and C. W. Palow. 2013. "Modern Safeguards for Modern Surveillance: An Analysis of Innovations in Communications Surveillance Techniques". *Ohio State Law Journal*, 74(6), 1071–1104.

Ishay, M. R. 2008. *The History of Human Rights: From Ancient Times to the Globalization Era*. Berkeley, CA: University of California Press.

La Rue, F. 2011. "Report of the [UN] Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression." Accessed at http://www2.ohchr.org/english/bodies/hrcouncil/docs/17session/A.HRC.17.27_en.pdf on 7 January 2019.

Marshall, T. and T. Bottomore. 1992. *Citizenship and Social Class*. London: Pluto Press.

Mitsilegas, V. and N. Vavoula. 2017. "The Normalization of surveillance of movement in an era of reinforcing privacy standards" in Philippe Bourbeau (ed) *Handbook on Migration and Security*. 232–251. Cheltenham: Edward Elgar.

Sony Pictures Home Entertainment. 2007. *The Lives of Others*.
Thomas, T. and T. Bottomore. 1992. *Citizenship and Social Class*. Pluto Press.
UNWTO. 2018. *Compendium of Tourism Statistics, 2012–2016*. Accessed at http://statistics.unwto.org/content/compendium-tourism-statistics on 7 January 2019.
White House Press Office. 2014. "Remarks by the President on Review of Signals Intelligence". Accessed at www.whitehouse.gov/the-press-office/2014/01/17/remarks-president-review-signals-intelligence on 20 February 2014.
Wright, D., S. Gutwirth, M. Friedewald, P. De Hert, M. Langheinrich, and A. Moscibroda. 2009. "Privacy, Trust and Policymaking: Challenges and responses". *Computer Law & Security Review*, 25(1): 69–83.
Wilsher, D. 2012. *Immigration Detention: Law, History, Politics*. Cambridge: Cambridge University Press.

INDEX

Locators in *italics* refer to figures.