

CIGI Papers No. 186 – August 2018

# Open Data Endgame Countering the Digital Consensus

Bianca Wylie





---

Centre for International  
Governance Innovation

CIGI Papers No. 186 – August 2018

# Open Data Endgame Countering the Digital Consensus

Bianca Wylie

---

## CIGI Masthead

### Executive

President **Rohinton P. Medhora**  
Deputy Director, International Intellectual Property Law and Innovation **Bassem Awad**  
Chief Financial Officer and Director of Operations **Shelley Boettger**  
Director of the Global Economy Program **Robert Fay**  
Director of the International Law Research Program **Oonagh Fitzgerald**  
Director of the Global Security & Politics Program **Fen Osler Hampson**  
Director of Human Resources **Laura Kacur**  
Deputy Director, International Environmental Law **Silvia Maciunas**  
Deputy Director, International Economic Law **Hugo Perezcano Diaz**  
Director, Evaluation and Partnerships **Erica Shaw**  
Managing Director and General Counsel **Aaron Shull**  
Director of Communications and Digital Media **Spencer Tripp**

### Publications

Publisher **Carol Bonnett**  
Senior Publications Editor **Jennifer Goyder**  
Publications Editor **Susan Bubak**  
Publications Editor **Patricia Holmes**  
Publications Editor **Nicole Langlois**  
Publications Editor **Lynn Schellenberg**  
Graphic Designer **Melodie Wakefield**

For publications enquiries, please contact [publications@cigionline.org](mailto:publications@cigionline.org).

### Communications

For media enquiries, please contact [communications@cigionline.org](mailto:communications@cigionline.org).

🐦 [@cigionline](https://twitter.com/cigionline)

Copyright © 2018 by the Centre for International Governance Innovation

The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the Centre for International Governance Innovation or its Board of Directors.



This work is licensed under a Creative Commons Attribution – Non-commercial – No Derivatives License. To view this license, visit ([www.creativecommons.org/licenses/by-nc-nd/3.0/](http://www.creativecommons.org/licenses/by-nc-nd/3.0/)). For re-use or distribution, please include this copyright notice.

Printed in Canada on paper containing 100% post-consumer fibre and certified by the Forest Stewardship Council® and the Sustainable Forestry Initiative.

Centre for International Governance Innovation and CIGI are registered trademarks.

---

## Centre for International Governance Innovation

67 Erb Street West  
Waterloo, ON, Canada N2L 6C2  
[www.cigionline.org](http://www.cigionline.org)

---

# Table of Contents

vi	About the Author
vii	About the Global Economy Program
1	Executive Summary
2	Introduction
3	A Brief History of Open Data, Open Government and the Open Movement
5	The Digital Policy Landscape
8	Exploring Possible Challenges with Open Human Behavioural Data
15	Policy Guidance
15	Conclusion
16	Works Cited
22	About CIGI
22	À propos du CIGI

---

## About the Author

**Bianca Wylie** is a CIGI senior fellow. Her main areas of interest are procurement and public sector technology. Beyond her role at CIGI, Bianca leads work on public sector technology policy for Canada at Dgen Network and is the co-founder of Tech Reset Canada. Her work at CIGI focuses on examining Canadian data and technology policy decisions and their alignment with democratically informed policy and consumer protection.

Bianca has a dual professional background in technology and public engagement, having worked at both Thomson Reuters and Swerhun Facilitation. She is a strong advocate for open government. Bianca founded the Open Data Institute Toronto (2014) and co-founded Civic Tech Toronto (2016). In 2017, Bianca took part in a City Hall Task Force organized by the University of Toronto School of Public Policy and Governance and she sat on the Ontario International Open Data Charter Advisory Board. Bianca is a guest lecturer, a tech columnist and a speaker on open government and public sector technology policy. Bianca currently sits on the Toronto Public Library's Innovation Council. She holds a bachelor of arts in political science from York University.

---

# About the Global Economy Program

Addressing limitations in the ways nations tackle shared economic challenges, the Global Economy Program at CIGI strives to inform and guide policy debates through world-leading research and sustained stakeholder engagement.

With experts from academia, national agencies, international institutions and the private sector, the Global Economy Program supports research in the following areas: management of severe sovereign debt crises; central banking and international financial regulation; China's role in the global economy; governance and policies of the Bretton Woods institutions; the Group of Twenty; global, plurilateral and regional trade agreements; and financing sustainable development. Each year, the Global Economy Program hosts, co-hosts and participates in many events worldwide, working with trusted international partners, which allows the program to disseminate policy recommendations to an international audience of policy makers.

Through its research, collaboration and publications, the Global Economy Program informs decision makers, fosters dialogue and debate on policy-relevant ideas and strengthens multilateral responses to the most pressing international governance issues.





---

## Executive Summary

The state exerts immense power through the ways it uses technology and data. This includes its role as a data steward, overseeing ever-growing amounts of information about government operations and people. As the narrative about data being the core of the future economy ramps up, so will the perceived value of government-held data. The idea of the state sharing some of its power and assets by sharing data is central to the open government data movement whereby governments publish data that anyone can freely access, use and share (Open Data Institute 2017).

In the broad set of interrelated governance and regulatory challenges currently brewing in the global data ecosystem, open data is but one small piece. Others include: power asymmetry between commercial actors in the market; a lack of informed consent models for data collection both online and in public spaces; and legislative frameworks that do not provide consideration or process for collective data ownership, mass consent and privacy as a public good. Given the range of unresolved challenges related to modernizing data policy and law (Scassa, forthcoming 2018; Ciuriak 2018; European Commission 2017; Fairfield and Engel 2015), there is a risk that current open data policies are creating loopholes, which may have unintended consequences.

We are currently facing large ethical and moral questions related to keeping options available for an untracked life. There is a heightened need to be intentional about maintaining elements of human existence that are unknown, unsellable and unable to be used commercially or prejudicially by anyone. Given this context, government-held data related to humans and human behaviour, such as health data, data about usage of urban spaces and facilities, education data, data about usage of social services programs and many more, may warrant specific and immediate attention from an open data policy perspective. Data about people, even when anonymous or aggregate, needs a special approach that may be hard to reconcile with openness. Without such an approach, the packaging, publication and use of human behavioural data as open data could potentially further commoditize humans and human behaviour. Specifically, it could: expand surveillance capitalism in public spaces; commercialize,

financialize and further outsource public service planning and delivery; and exacerbate tensions in the construct of liberal democracy.

These three technology trends are well known and have been documented for years by scholars, journalists and civil society. The issue raised in this paper is precisely a question about how these issues intersect with the idea of open data. Are open government data programs positioned to add fuel to these fires by adding human behavioural data to the market as open data? The current use of commercially collected human behavioural data is not under control from a legislative and policy perspective. Is it worth adding more complexity to the landscape by publishing additional open human behavioural data from the government side? What are the trade-offs to consider in doing so?

It appears to be time to pause the open data program and assess its impacts and role in a much-needed public conversation about data and democracy. Ongoing developments in technology should not continue to impact society by accident or state negligence. We should not be further commodifying our existence through a government program that has not attained broad social licence. While some may think that any demand for regulation around data must come from residents, history has shown that technology has been unhindered in recent decades due to a fundamental lack of understanding about how it works and its societal impacts (Zuboff 2015). Thus, the state must be the proactive actor in asserting its role and place in the conversation. The policy consequences of data and technology require more thorough democratic engagement, including the questioning of current social norms that were never informed by broad social consent. Part of this conversation includes the role of open data in our democracy. Simply put: what is the open data endgame?

---

## Introduction

Don't eat capitalism raw.

— Shoshana Zuboff (2014) on Thomas Piketty's *Capital in the Twenty-First Century*

The way in which the state takes on a moral role in society is not fixed; it is heavily influenced by history and culture and is persistently reshaped through democratic participation. Beyond regulation, a suite of government policies, from economic development to infrastructure, inform and influence the defined or blurred line between market and state, including the values that are exerted in these policies. Open government data is a policy issue that sits in this context and straddles the dividing line between market and state. Consider open government data as data that the government publishes for anyone else to use. The core idea explored in this paper is whether human behavioural data may warrant exemption, for the time being, from publication by the government through its open data programs due to the potential for this data to be used to further commodify human behaviour.

### What Is Human Behavioural Data?

For the purpose of this paper, human behavioural data is defined as both aggregate and de-identified data about people. Aggregate human behavioural data is the data of patterns: how and which people use a park, where people ride their bikes, education results by school and so on. It is not data about one person, but data about a group of people. There is also de-identified data, where humans are made anonymous by changing identifiable elements of their data. This means individual-level data can be released, but without knowledge of who the data is about. These are two distinct cases, but for the sake of simplicity they will both be lumped into the broader category of human behavioural data. While it is not simple, nor in some cases even possible, to cleanly separate human behavioural data from many other types of data, the idea is to create a specific way of thinking about data that is generated by, or identifies, humans and their general traits and/or behaviours in a data set.

### Blurred Line Between Market and State

“This amoral character of economic life amounts to a public danger,” wrote Émile Durkheim, a founding sociologist of the early twentieth century, in his work *Professional Ethics and Civic Morals* (1957). It reflects an idea that the state has a vital role to play as a moral arbiter in society. A modern day exploration of the implications of Durkheim's ideas might involve a discussion of the state's role in consumer protection in a liberal democracy: a study of the state's ethical requirement to participate in the organizing of the market, to manage its influence and impact on human life.

Before focusing on the potential risks of publishing human behavioural data as open data, some high-level context regarding key concepts will be provided. This includes a brief history of open government data policy, both globally and in Canada, privacy norms in Canada and thoughts about the categorization of data. Then some trade-offs to consider in the context of open data policy will be explored, given the current and potential uses of human behavioural data. The paper will close with several suggestions on policy approaches to consider in the management of these and other related unknowns in the open government data policy space. There is a wide breadth of content raised here, and this is an early set of thoughts shared to help support and continue the discussion about open data and its future. With this proviso, and explicit self-awareness, readers are requested to think of this paper as part of a conversation about issues to consider in the evolution of open data policy, knowing that additional and narrowed analyses and research are required and that many elements of this work have been touched on many times for many years now in research regarding data, privacy and regulation.

### Defining Open Government Data

Open data has three core qualities: that it is free (it costs nothing to obtain), that it is made freely available for use (there are no restrictions on how it can be used) and that it is machine-readable (Open Data Institute 2017). Open data can be published by anyone with the capacity to generate data and host it online. Open data publishers can be private or public companies, governments, residents, civil society institutions or others. This paper focuses on open government data and related policy. A few examples of common open government data

are transit data, budget data, environmental data, research data, geospatial data and many more. Geospatial data — data indicating the physical location of a range of assets and infrastructure — comprises a large amount of open government data.

One of the defining qualities of open data is an option to license it for use without dictating who can use it and how it can be used. When licensed in this way, as much open government data is, the openness means it is made available to anyone that has the capacity to find and use it. Herein lies a double-edged sword, in that while the publication of open data is intended to democratize its use, the number of actors that can truly make use of it is small, as they require infrastructure to access the data, highly technical skills, access to technical assets and, sometimes, intellectual property and capital. As such, often these are established institutional and corporate actors, not residents (Argast and Zvyagintseva 2016; Weerakkody et al. 2017; Robinson and Ward-Mather 2017).

---

## A Brief History of Open Data, Open Government and the Open Movement

### History of Open Data Globally

Open data does not have one distinct origin. In a broad historical sense, its foundational roots could be linked back to open science in the late sixteenth and early seventeenth centuries (David 2008). In more recent history, it is tied to the notion of transparency as essential to the rule of law and good governance as per Louis Brandeis, associate justice of the US Supreme Court (1916–1939) (Berger 2009), the philosophical notion of an open society (Popper 1945), as well as to calls for open access journal publishing by the US Office of Scientific Research and Development in 1945 (Wolpert 2013). These are a few schools of thought that came before the internet era of the 1990s, which is the technological development that enabled the open data movement to exist.

The open government data movement started in the United States and the United Kingdom, with the US open government data portal, data.gov, launching in 2009 and the UK version, data.gov.uk, launching in 2010 (Janssen 2012). The open government data movement came of age as part of a broader openness agenda, called open government. In 2011, eight founding signatories (Brazil, Indonesia, Mexico, Norway, the Philippines, South Africa, the United Kingdom and the United States) endorsed the Open Government Declaration and announced their first national open government action plans. Since 2011, the Open Government Partnership has grown to 75 member governments that write biannual plans comprised of state commitments to increase openness in their governance approaches.

Open government data as a policy direction has four key drivers: transparency, economic development and innovation, participatory governance and public sector use (Janssen 2012; Sieber and Johnson 2015). The open data movement has been slow in practical uptake within government, but even with slow uptake, government has considerably outpaced the private sector, the non-profit sector, the academe and others, in its commitment to the publication of open data. Although innovation is said to be a driving rationale for sharing data, this narrative comes predominantly from government in terms of how it can support innovation and economic development, by sharing data as an input for business development. The private sector is a less frequent contributor to the open data movement. Companies generally hoard their data, which in turn gives them power over their stakeholders (Ciuriak 2018).

In many countries, open data was heralded by civil society as an opportunity to push back against corruption through demanding transparency in data publishing. In others, open mapping data made humanitarian interventions possible in the wake of natural disasters, helping aid workers navigate damaged infrastructure with real-time mapping updates provided by residents and international crisis mappers (Brandusescu, Sieber and Jochems 2015; Landry et al. 2017).

A range of present-day work on global open data can be explored through two initiatives: the Open Government Partnership and the International Open Data Charter. Beyond these two formal organizations, some governments have adopted an “open-by-default” model. In Ontario, for

example, all government data should be made open unless it is exempt for legal, privacy, security, confidentiality or commercially sensitive reasons.<sup>1</sup>

The impacts of open government data programs have been the subject of much study, none of it very conclusive. There are a variety of questions regarding whether the right metrics are even being considered in assessing the programs (Zuiderwijk, Shinde and Janssen 2018). Common metrics for open data relate to what has been published and how, rather than studies of the way data has been used, which are challenging to conduct. Suffice it to say that open data cannot be termed a success nor a failure to date, due to the immense array of activities it is intended to support, differing levels of government commitment to the programs and challenges in measuring the breadth of programs it involves (Longo 2011; Sieber and Johnson 2015; Robinson and Graser 2016).

## History of Open Data in Canada

Open government and open data were adopted as policy under former Prime Minister Stephen Harper's Conservative government in 2012. The Canadian approach put economic development and innovation on centre stage as the rationale for open government data programs, with open data heralded as a new resource to be exploited. This enthusiasm for business development through open data was bolstered by McKinsey's infamous 2013 report that put the value generation of activities related to global open data at US\$3 trillion (Manyika et al. 2013). Beyond the economic development mandate, the Canadian state eagerly tapped into the language of improved democracy and scientific progress through open data, just two years after repealing the mandatory long-form census — an early flag of how easily the initiative can be used as a political communications device. The Canadian open data narrative has continued to evolve with a change of government and is now organized around policy co-design and civic engagement as well as economic development. The co-mingling of civic and business rationale for openness has created a wide range of stakeholders involved in the open data community, from entrepreneurs to activists to journalists, similar to research findings describing the open data stakeholder community in the United Kingdom (Bates 2012).

The Canadian open government and open data program is progressing, although slowly. There are a range of implementation issues, from challenges in extracting data due to legacy information technology (IT) systems to cultural resistance toward openness in communication about government operations for fear of reprisal and negative press. These issues are widespread globally and have led to the creation of the term “open-washing.” Open-washing describes the goodwill that politicians accrue by declaring that their government is open, while in reality having a degrading performance on access to information programs and data requests from civil society actors. “Open-wishing” is another term frequently used to describe open government programs. This phrase describes an open data approach whereby data is released through a government website, often called a data portal, without much additional program support, followed by political announcements about how societal change will follow naturally from this action (Howard 2016; Ruijter, Grimmelikhuijsen and Meijer 2017). These two issues are visible within Canada's open data program. All levels of Canadian government should work to address these issues. Many of the stated aims of the open data program have merit, but they require intentional support, funding and cultural evolution, including changes to roles and responsibilities within the public service.

While the open data movement has mixed results on achieving its stated aims, it has also opened the door to a set of unintended (or at least underdiscussed) policy consequences. These consequences relate to the potential release of human behavioural data into the market — data that was originally provided to the state by its residents with an understanding that the data would be managed by the state in return for program and service delivery, a clearly understood element of the social contract (Wylie 2018b).

---

<sup>1</sup> See [www.ontario.ca/page/ontarios-open-data-directive](http://www.ontario.ca/page/ontarios-open-data-directive).



---

# The Digital Policy Landscape

Western governments, North America in particular, have enabled a legislative and policy vacuum regarding consumer protection and technology products, in particular in the context of data products (Wylie 2018a). Lacking clear guidelines as to what sort of products and services are permissible, technology companies have pushed boundaries and created norms in domains that few people understand, in particular around the use of their data. People often share their data without a full picture of how it is being used (Lessig 2000; Zuboff 2015). Moral and ethical decisions have been made about the shape of our current society through the use of these products without adequate public discourse about their impact or any broad and informed social consent. When Google began to drive cars around to take photos of streets, houses and infrastructure, there was no public conversation about the implications of this activity. It was allowed, in part, as there was no forum for the discussion (Zuboff 2015).

Technology companies have also been enabled by a light-touch regulatory environment that supports the idea of “permissionless innovation” as a guiding rationale to maximize economic growth. Permissionless innovation is an idea that supports innovation as a default for business, in contrast to the usual government use of a precautionary principle that emphasizes a need for business to prove that their actions will do no harm to people, as consumers (Thierer 2014). While some might argue that the precautionary principle may be too constraining, or lead to a totalitarian construct, there are already industries, such as the pharmaceutical industry, that are subject to immense amounts of regulatory process leading up to any product release. It is not without precedent that government has imposed regulations without dampening market interest in participating in the economy. One of the more difficult things to assess with data and technology, however, is how to define harm, both at the individual and societal level.

These two trends, a lack of up-to-date consumer protection policy for data use and the permissionless innovation economic development narrative, have resulted in an increase in, and the normalization of, the commodification of human

behaviour — a phenomenon sometimes referred to as surveillance capitalism (Zuboff 2015). The lack of public knowledge about how people’s data is being gathered, used and sold has been made clear by exposés regarding the operations of firms such as Facebook, Uber and Palantir (Solon 2018; Frizell 2014; Waldman, Chapman and Robertson 2018). Updating existing privacy laws to keep pace with the evolving technology industry is an ongoing challenge that exacerbates the negative impacts of these issues. The model of treating data as a private good is also problematic. Taken as a whole, the impacts of technology, in terms of how data and data products have been and could be impacting people’s quality of life, present many unknowns for residents, policy makers and the legal community.

In the meantime, over the course of the past decade, open data programs have become entrenched in global government operations with varying degrees of support and authority. Regardless of how well these programs have performed, they have attained a mostly normative status of being inherently good and pro-democratic (Sieber and Johnson 2015), although several issues have been raised to challenge this assertion (Longo 2011; Bates 2012; Johnson 2014; Johnson et al. 2017).

While open data has attained a relatively innocuous status as a government program, there may be new pressure to increase the amount of open data published by government due to state interest in developing technology talent and competitive advantage in machine learning and artificial intelligence (AI) (Villani 2018). France and the United Kingdom have recently made significant commitments of political and financial capital to state AI programs (Thompson 2018; Kahn and Morales 2018). Adding to the potential increase in open government data publication, there will be new unknowns related to the use of public and private data for machine learning and AI (Thereaux 2017). While the utility of aggregate and de-identified open data for machine learning needs additional analysis, as it is not the ideal granular data for the activity, the range of ways it could be used sits in a space full of unknowns that are constantly evolving as computational gains are made. In response to the unknowns posed by machine learning and AI applications, several countries, including the United Kingdom, Canada, France and the United States, are investing in research institutions to explore the ethical

application of AI (Thereaux 2017),<sup>2</sup> with Henry Kissinger recently calling on the United States to urgently invest in this effort (Kissinger 2018).

The assumption implicit in these programs is that government is a neutral arbiter, suitable to perform an ethical analysis of AI. While they certainly should be making the effort, it is also important to consider independent analysis that comes from a range of communities and civil society regarding their data and its use. There is evidence that many residents are increasingly marginalized by the ways their data is used within the state by simple rules-based software systems, whether from a privacy perspective, a discrimination perspective or both (Citron 2007; Eubanks 2017). Put another way, in a critical assessment of open government policy (Bates 2012), data and open data, and the software and AI they feed, may well serve to further “empower the empowered” and marginalize the marginalized (Gurstein 2011; Bates 2012; Eubanks 2017; Longo et al. 2017).

## Human Behavioural Data: An Open Data Exception for Consideration

At the heart of this paper is a suggestion that in open government data programs, human behavioural data should be considered differently from other types of data. There is an urgent need to define collective privacy and mass consent and to challenge the idea that aggregation or anonymization resolve all of the issues with open or shared data because they do not. Neither approach provides answers to questions about keeping things about ourselves as humans and as collectives unknown and safe from commodification. Nor do these approaches address control and use — keeping certain data available only to the state or to non-commercial entities for use in the construct of a well-understood social contract. These are issues to be addressed through new policy and regulation around data collection rules, data licensing, data minimization and data deletion, as a start.

The implications of the collection, use and sharing of data that comes from, or is tied to, humans presents different challenges than many other types of data (Zuboff 2015). Currently, much of the behavioural data that is discussed in the

mainstream media is collected in a commercial capacity, at a personal level, through the use of products and services such as Facebook and Google, or mobile phones or fitness trackers. Less discussed, but possibly more important, in terms of the evolution of the roles of market and state, is the data collected and held by governments. The root of government’s ethical commitment to manage data is to support human well-being. Whether they do it or not is a political decision. Regardless, this is never the primary role of industry.

Current government open data policies were primarily designed for budget data, program data, environmental and geospatial data, and asset-based data — they were not designed for human behavioural data. The policies have safeguards to protect the privacy of the individual but not the collective. Governments have not received broad social consent to open the behavioural data they hold. Much of the data held by governments was collected before the idea of open data policies or being open by default even existed.

Again, most of the data that is, or could be, released under the open government data program is fine to continue with and expand on. This paper is not suggesting the pace of the open data movement should be slowed where it is working. In cases where the data is human behavioural data, a new approach may be needed. This may mean that, out of caution, some open data should not be published as such — this does not mean it cannot be shared with other stakeholders under certain conditions. Open government advocates may challenge this approach for three main reasons.

The first is that open data creates accountability within government. This is true, and so long as data regarding programs and costs and governmental operations are all published, this would continue on with little impact. Also, the accountability piece related to government data is often more within the domain of freedom of information regimes than open data. Second, some argue that open data belongs to residents as they are taxpayers and thus funded its creation. This is a somewhat libertarian argument that places the program’s rationale in market terms and misses much of the context of the data and its relation to government, governance and the social contract. Third, open data supports the advancement of science, assuming that sharing data spurs innovation both in research and economic

---

<sup>2</sup> See [www.cifar.ca/assets/pan-canadian-artificial-intelligence-strategy-overview/](http://www.cifar.ca/assets/pan-canadian-artificial-intelligence-strategy-overview/).

development. While true in theory, the practical impacts of these programs require further study.

All of these rationales for open government data programs require further consideration, as policy about open human behavioural data should be informed by a conversation about trade-offs, not a binary right or wrong approach. There is a large amount of government data that is not related to humans — this should still be released if it qualifies under the current program. Again, this suggestion around the particular consideration of human behavioural data is not intended to slow down the pre-existing beneficial cases of open data publishing; much of that data is not related to people. However, in regard to human behavioural data specifically, it is defensible that the desires and impacts related to the people in the data, both as individuals and collectively, must be assessed prior to carrying on with an open-by-default mandate.

## Privacy as a Public Good and Collective Privacy Rights

The current way human behavioural data is treated from a privacy and rights management perspective is problematic, even in the more progressive iterations of data protection law such as the European Union's General Data Protection Regulation. Privacy laws are organized around the idea of personal data and singular individual data. The concept of consideration of our collective societal rights around data is undefined in Canadian law. Recent works on this issue suggest the idea of privacy as a public good (Kwecka et al. 2014; Fairfield and Engel 2015).

If we are to be intentional about changing how data is organized for the public good, the time to exert that intention is now. This role will fall to the state, not the market. Many commercial actors in the data space serve to benefit from an asymmetry of power and desire more detailed data, not less of it (Ciuriak 2018). As for a third way, such as community data management, it is generally unrealistic to consider the full range of data governance as something that all individuals would or could manage themselves (Obar 2015). Those that have the capacity and desire to do this work would likely not be representative of the general population, which opens up another avenue for the further marginalization of non-participating communities.

Existing practices regarding human data collection in public spaces need to be reconsidered as well

(Robinson 2018). As public awareness catches up regarding the current implementation of surveillance capitalism, it might be expected that a social backlash will create political demand to design stronger consumer protection around data, legislative updates and the exploration of collective privacy — privacy as a public good (Fairfield and Engel 2015; Marshall 2018). With this may come the rejection of the escalation of surveillance capitalism in public spaces and calls for dismantling some of its current architecture and infrastructure (Bailey and Caidi 2005; Zuboff 2015).

The persistent narrative that privacy is dead, born of Eric Schmidt (Esguerra 2009), is untrue — there are vast amounts of data that are yet uncollected in public spaces and others that are solely in the domain of the government or independent third parties with whom individuals have chosen to share data. The notion that the development and deployment of rapidly advancing information technologies are an inexorable force against which governments and other forms of social authority are impotent, often referred to as technological determinism (Kline 2015), has infected public discourse, dampening serious discussions about technological governance, to the pleasure of multinational corporations (Zuboff 2015; Pasquale 2018). While the state can be a malicious actor regarding the abuse of data, it is also the actor that can create laws to roll back existing data collection practices and their legality, if necessary. Data ownership must be understood as a critical foundational piece in any discussions of data privacy, control and use (Scassa, forthcoming 2018).

We are socially under-equipped to have conversations about data ethics and their political economy. We are struggling to understand who must lead in creating space for these conversations. We must address these challenges and explore how to involve questions about the morals and ethics of technology in a series of public consultation and public engagement activities. In the absence of this activity, techno-determinism and surveillance culture will continue to advance unchecked (Lyon 2017). Avenues to define and claim collective privacy rights, information rights, social norms around data collection and thoughts around continuing efforts in digital constitutionalism, can continue to be identified and opened as channels for public discourse (Caidi and Ross 2005; Gill, Redeker and Gasser 2015; Taylor 2017).

## Community Data Protection: Mass and Ongoing Consent

Finally, before getting into the discussion regarding the challenges posed by open human behavioural data, it is important to keep in mind that existing laws inform and override any open data policy. But this is where the problem becomes nested. The foundational law for privacy regarding government-held data in Canada, the Privacy Act, is built around the idea of individual data, not aggregate or de-identified data. It does not include any notion of collective data rights, privacy as a public good or an updated informed consent path required of a person to allow their historical data to be released, even in a de-identified and aggregate way, as open data (Scassa 2014; Sánchez and Viejo 2017). So, while the policies and data charters are one layer, the underlying legal direction is out of date for the cases that will be discussed here. In the absence of explicit protection of our aggregate and de-identified information, it could be argued that a lot of historical human behavioural data could be released as open data. Statistics Canada already releases aggregate data as open data: it can do so because aggregate data is not personal information within the statutory definition of the term.

---

## Exploring Possible Challenges with Open Human Behavioural Data

Having set the context regarding the history of open data, the reasons the program may accelerate, the baseline context of privacy for data in Canada and the idea of open human behavioural data, the paper will now turn to explore three potential risks related to the publication of open human behavioural data and how it might feed into the further commodification of human behaviour. Prior to launching into this evaluation, please remember, the question is not whether these issues related to data are occurring — they are — it is whether open government data programs could potentially accelerate these issues. If so, to what degree, and what are the next policy steps to take to address the trade-offs inherent in the ongoing release of open government data?

## Argument 1: The Individual

### Behavioural Manipulation

Privacy is not dead for the privileged. In Canada, data profiling is facilitated by a range of data sources about individuals and accelerated by the use of various consumer technologies (such as social media platforms, fitness devices and mobile phones).

The Privacy Act is the federal legislation that oversees the use of data held by the state to deliver public services. The core guiding principle of this act is that data can only be used for the purpose it is collected for, meaning there is not one large state database that profiles residents through all of their various sets of data. The Personal Information Protection and Electronic Documents Act (PIPEDA) is the federal legislation that oversees data collected and used by the private sector.

Public space in the smart city presents a legislative grey zone between the two. As commercial interests continue efforts to assert ownership over data collection in public space in the smart city context (Kitchin and Lauriault 2014; Edwards 2016), the remaining public spaces that are arguably safe from tracking are threatened. This leads to the loss of reasonable expectations of privacy, accelerated surveillance and behavioural influence (Murakami Wood 2013; Zuboff 2015; Edwards 2016; Mattern 2018).

Various corporations continue to make efforts to participate in public space data collection as smart city vendors (Murakami Wood 2015; Lauriault, Bloom and Landry 2018). They are starting to align themselves, in word, with the open data movement. As will be argued here and later in this paper, this is fundamentally problematic in that it assumes corporate ownership of data collected in public space, through methodologies that may be in conflict with informed consent. In public space, the default assumption should be government (public) ownership. The idea that corporate ownership of urban data has taken root shows how new norms can quickly emerge within a policy vacuum when there is no precedent. The move by corporate actors to support open data is strategic in that it aligns private sector intentions with an openness approach already endorsed as a normative good with government, but takes the policy out of context and abuses



the lack of broad informed social consent to proceed with this application in public space.

The tracking and storage of data related to fine-grained personal preferences and behaviours, both online and offline, through the use of social media, telecommunications, fitness accessories, smart home devices and more has caused a recent explosion in behavioural data collection. The firms or data brokers with access to this volume of behavioural data now seek to move beyond the satisfaction of stated user preferences (such as delivering relevant results in a search engine) to the prediction of those user preferences. As Google's chief economist Hal Varian (2014) put it, "Google should know what you want and tell it to you before you ask the question." Achieving this goal, which generally occurs in a low-stakes retail environment, can and does provide users with a range of services with considerable convenience. However, there is also reason to believe that this predictive power could create undesirable consequences through subtle and pervasive manipulation of behaviour, as seen in the recent case of Cambridge Analytica's use of Facebook data. In her account of surveillance capitalism, Shoshana Zuboff (2015) argues, "in Varian's economy, authority is supplanted by technique, what I have called 'the material dimension of power,' in which impersonal systems of discipline and control produce certain knowledge of human behaviour independent of consent." There is a case to be made for placing restraints on access to the behavioural data that enables this form of influence, including the capture of human behaviour in public space. This, much like the categorization of open human behavioural data, is not a clear or distinct set of data, but again, it is a concept to consider in the framework of updating consumer protection regulation (Marshall 2018).

Insights from the realm of behavioural economics have demonstrated that human decision making is often a product of unconscious heuristics and shortcuts rather than conscious, willful deliberation. Human minds hold many biases and irrational features, causing behaviour that is inconsistent with assumptions from classical economics around rationality and autonomy (Tufekci 2014). Experimentation in behavioural domains has demonstrated that people will respond differently to the same question depending on how it is worded, such as whether the consequences are described in terms of losses

or gains foregone, or which option is listed as the status quo (Sunstein and Thaler 2005). Therefore, the actor responsible for arranging the context in which individuals make decisions, the "choice architect," is able to exert subtle but powerful influence on the decisions people make (Tufekci 2014). Consider the impacts of this power in a smart city context.

Karen Yeung (2016) argues that the ability to subtly manipulate behaviour without resorting to what would technically count as coercion would be magnified by big data, leading her to dub this form of behavioural manipulation a "hypernudge." Data collected through digital platforms and the Internet of Things makes it possible for choice architects to construct personally tailored choice contexts for each subject. Instead of "nudging" the behaviour of each affected individual in the same manner, access to human behavioural data allows the execution of much more intricate feats of manipulation, influencing the behaviour of different individuals in different directions. Furthermore, the continuous feedback of new data allows the real-time restructuring of individual choice contexts based on their past decisions (ibid.).

While big data allows the identification of hitherto hidden correlations, experimentation is required to establish the direction of causality. As Zuboff (2015) articulates in her account of surveillance capitalism, if a choice architect wishes to influence individual behaviour, it would need to engage in experiments on the individuals from whom its data is collected. Witness, for instance, the experimentation carried out by Facebook, wherein the sorting algorithms that determine the content visible to users on their homepage was altered for a sample of users, in order to determine if Facebook could change its users' apparent moods (Meyer 2014). The degree to which this past behaviour is a guarantee of future behaviour is full of unknowns. This is not to say that with all of this data there is a guaranteed opportunity for control, or a known degree to which it can be exerted.

Furthermore, the unprecedented volumes of data now being collected by the likes of Facebook, Google and Amazon through their platforms, as well as the ongoing implementation of the Internet of Things, enables the collection of personal behavioural data to extend into the physical world, in a variety of ways both explicit and indirect (Murakami Wood and Day 2018). This could grant big tech companies access to new tranches of data to add to their repositories, expanding

data profiles that are already far more advanced than the profiles governments hold on the large majority of residents (Yeung 2016). The public policy modifications that have been the focus of nudge research thus far are carried out by elected governments ultimately accountable to the general populace. The major players in the domain of big data are corporations whose prime directive is the pursuit of profit. While it may be possible to justify some nudges in public policy as a means of promoting user welfare, the fact that the agents most capable of executing this form of behavioural influence are profit-maximizing corporations renders this line of reasoning more dubious.

Zuboff's (2015) notion of surveillance capitalism illustrates the potentially harmful consequences of these tools of behavioural prediction and influence. Surveillance capitalism may be a misunderstood term by many, as the word surveillance is often immediately assigned to policing and privacy invasions. While this piece is true of an economy that is ever-increasingly tracking and monitoring citizens by way of data and sensors, including cell phones, it is the economy of this activity that is the new paradigm. Commoditizing human activity and understanding preferences to define human behaviour, in everything from search results to commerce to physical activity, has been on the rise since the early days of the internet. What is happening today, however, with the rise of smart cities, is the possibility for data collected in public spaces, perhaps owned and licensed by corporations, to be added to the inputs for this commoditization.

This is a next step in surveillance capitalism in that it may be enabled by liberal democratic states in the rush to apply smart city technologies to public spaces. With vendors claiming that they will open data, tech companies want to be the owners, collectors and stewards of data, but with little clarity available to the public on how they may want to use this data in their operations and product and service offerings. Should they be offered public space by the state to do this experimentation and business because they are willing to commit to open data? The commodification of human behaviour is old; this is an acceleration, not a new activity (Murakami Wood and Ball 2013). The question is whether this acceleration should be aided by enabling the collection and/or

publication of open human behavioural data, by corporations and governments alike.

## Argument 2: The Nation

### Commercializing and Outsourcing the Public Service

“It seems to me that the real political task in a society such as ours is to criticize the workings of institutions, which appear to be both neutral and independent; to criticize and attack them in such a manner that the political violence which has always exercised itself obscurely through them will be unmade, so that one can fight against them” (Chomsky and Foucault 2006).

The public service, in North America in particular, is fundamentally disempowered to apply technology to its mandate across a wide range of operational areas. It takes a certain amount of institutional negligence to arrive at this point and whether this was political malfeasance or bureaucratic sclerosis by senior management or unethical business practice by government vendors warrants its own independent study. Regardless of the cause, it is a departure in the long arc of history of the power held by the modern bureaucracy. Far from German political economist Max Weber's “iron cage,” the public service is in a cage of its own, constrained by a lack of technological investment, the outsourcing and privatizing of technology functions and vendor lock-in (Maley 2004). The implosion of the Canadian federal government's Phoenix payroll system is a case in point. A large project was contracted out to a vendor; the project went into major cost overruns, while unionized public sector IT staff — some of the very people the system was supposed to support — have been rendered powerless to engage in fixing the problem (Ireton 2018).

Power that was once increasingly condensed in the bureaucratic hierarchy is being shifted to the private sector through its design and control of technology systems that have long been central to government functions (Howlett and Migone 2014). Weber's theory of the bureaucracy saw one potential challenger to the expertise of public servants — small groups of private sector actors with specialized knowledge (Metz 2015). It can be argued that this challenger has arrived in the form of consulting and technology firms, with large technocratic influence being exerted on policy and procurement. There is an increasing amount of rhetoric and funding being directed

at technocratic improvements for governments, and a framing of these improvements as a normative good, a phenomenon that can be called the digital consensus. The digital consensus assumes the inherent good of any government IT modernization project, when in fact many are political decisions that often entrench the status quo, in addition to being both wasteful and counterproductive (Wylie 2017).

At this point, there is a risk that an expansive open data strategy could open the door to further erosion of technological capacity within government and increased privatization of government service delivery (Bates 2012; Johnson et al. 2017). Data is one of several core inputs to the planning of public services. The open government data program calls on government to continue opening vast amounts of data, data that to date has often only been managed or used by public servants — data that was also collected with no intention of ever being opened beyond government use. Add this historical data to new data from a set of environmental sensors, where more data will be available in real time, and the data can be productized and sold back to the state as a service (Johnson et al. 2017).

Take urban planning as an example. Through the creation of private digital infrastructure, whether provided for free or at a cost, new and better data is made available to both the public and the private sector. Rather than taking this data and using it as an input for public service planning, new products will be created and sold to government based on the product's ability to apply this data to planning contexts. This is a growing issue with the emerging area of data as infrastructure and the sensor systems and related architecture that will generate data about activities in public spaces. To ensure that the data that is generated is used in an accessible, transparent and democratically informed and policy-driven way, it is worth considering whether this data should be state owned. When it is state owned, data can be subjected to a critical lens regarding whether it should be made open or not, including a consideration based on whether the data includes information about human activity and an exploration of the possible positive and negative impacts of its release.

As discussed previously, the first rationale for this is ethical. Public servants have an ethical relationship to people's data and must use it within the ethical frame of government service

provision to make decisions. The private sector has no such responsibility in the use of this data. The challenge at the moment, precipitated by years of outsourcing technology and contract staffing in government, is that many technology companies have significant advantages over government in how they are able to perform data analysis and other transformational data services. This is true of data analysis and particularly true in regard to machine learning and AI. But when products use government data to support, for example, planning functions, whether for transportation planning, housing or other activities, there are going to be factors that need human awareness and professional expertise regarding the humans that said data represents. Then there is a new set of concerns related to the application of open data to analytics, machine learning and AI. Using data published in aggregate by the state or generated in public spaces by residents, the private sector will continue to offer new products and services that may be both opaque and problematic in how they apply data (Pasquale 2015; O'Neil 2016; Noble 2018). This is not to say that data use by the public sector guarantees responsible treatment of data — it does not (Eubanks 2017), but it does increase the potential to engage with how data is used due to a higher level of accessibility regarding the inputs of government services versus commercial products.

By buying products that run on resident-generated data instead of building capacity in government, a core layer of both infrastructure and service delivery could become privatized and entrenched in government budgets (Ranchordos and Klop, forthcoming 2018), much like the past 20 years of enterprise and custom software purchasing. In this case, rather than have the public service define the needs for data collection and use, the private sector sets the collection model. This is an important moment to pause and consider whether having a broadly open government and related open data model is truly about democratization or whether it is about commercialization and outsourcing. The fundamental relationship between people and their government is accountability. There is no accountability between a government and a vendor beyond a contract. There is no clear mechanism nor approach to be able to understand the business strategies and medium-terms goals of the private sector when it comes to creating and selling open data products to the state, although efforts have been made to assess the landscape (Zeleti, Ojo and Curry 2016). There are two approaches to take with

this issue: the first is to limit the sensitive data that can ever be published by governments about its residents, whereas the second one can work from the other end — procurement. This could include defining products and services that can only be created by the state, in order to maintain awareness and inputs on the full line of decisions, from the data involved, how complete it is (or not) and how it will be used in public service provision.

The current procurement models that government uses to buy technology can lead to the purchase of closed, proprietary products — software and systems that make use of complete data modelling that cannot be understood or parsed methodologically by its buyer, government and the people it is used on or for (Pasquale 2015; Eubanks 2017). There is also a mounting technical literacy gap between vendors and government, as well as all the issues inherent in code-based programs and dashboards being used in place of rules interpreted by humans (Citron 2007; Kitchin, Lauriault, and Mc Ardle 2015; Janssen and Kuk 2016; Eubanks 2017).

As discussed previously, the data literacy levels and financial resources available to individuals or civil society actors are miniscule compared to corporate power and capacity to carry out intense data analysis, including using data and open data as inputs to machine learning and AI. In this context, enabling access to shared open data stores is questionable in terms of how much it might be able to level the tables for residents or civil society. Some go as far as to say that open data as currently managed does not improve democratic outcomes (Gurstein 2011; Ruijter, Grimmelikhuijsen and Meijer 2017). Some argue that a move toward broad corporate open data publishing will exert more fairness in the market — shifting power from large companies to small (Tennison 2018). However, the sales feature of products sold to government as part of an open ecosystem may be a false claim to being a good actor. Opening data when you are the main actor commercially situated to benefit from doing so is not opening data in good faith.

The model of using government data, including open data, in commercial products that were once the domain of planners and civil servants accelerates the idea of government as a platform (O'Reilly 2009; Goldsmith and Kleiman 2017). There is likely erroneous thinking within industry that open data published by governments will provide a neutral and logical input for public service planning. This, again, is the digital consensus,

an assumption of technical neutrality, one that overvalues and misapplies the quantitative over the qualitative in decision making (Mattern 2017). Unless one is well acquainted with the humans and human circumstances from which the data comes, it is open to misunderstanding and misapplication. This is likely to happen with commercial products created for government. Public servants have an important role to play in setting the terms, boundaries and interpretations of data in policy creation. Here we can turn back to the data aggregation theory to consider how data collected under corporate rather than scientific modelling may not be the appropriate frame for inputs to public service delivery and policy generation (Kitchin and Lauriault 2014).

In addition to removing the ethical practitioner from the inputs of service delivery, consider the issues with data quality that could be missed through the release of open human behavioural data — there is a high likelihood that this will be older data and possibly, due to problematic data collection methods, potentially discriminatory data and incomplete data (Pasquale 2015; Scannell 2016; Weerakkody et al. 2017; Noble 2018). It is vital that a non-commercial lens be applied to the use of any human behavioural data collected in a government capacity. Public servants are positioned to both know the issues inherent to data that may be used in a data product and to have the ability to intervene in the usage of any data products or services that may use open data as an input.

## Argument 3: Global Systems

### Accelerated Erosion of Liberal Democracy

After exploring some of the possible impacts of releasing open human behavioural data at the individual and national levels, and how it could be used to commoditize human behaviour in those contexts, the final lens to consider is the global systems level. Global cities of the future are often described as being informed by and responsive to real-time and constant data collection (Kitchin 2013; Foroohar 2018). Beyond the public realm, private data collection through a range of personal devices, from cellular phones to ride-sharing programs to fitness trackers, is on the rise as well. Taken together, they have the potential to significantly impact several core features and models of liberal democracy. One of the much-lauded benefits of ubiquitous data collection is that it enables precise measurement to take



the place of estimates and averaging. Yet, to a certain extent, our social comfort with the current economic system has always been predicated on imprecision in ways not fully appreciated.

First, consider the insurance market and any products and services reliant on credit. As data and data brokerages help these firms understand the risks inherent to covering and servicing individuals, the baseline model of risk pooling and averaging is threatened. In the insurance model for various products, the collective group enables protection by having a shared risk. As insurers and lenders become increasingly able to profile and withhold products and services for individuals, well beyond the credit score, more people will likely be cut out of a range of services — from housing to transportation to other critical pieces of infrastructure for a decent life. The burden of this responsibility is due to fall to the state, which in many liberal democracies is already buckling under debt (Ferozohar 2018). At a simplistic level, there are limited reasons to add additional data as collected through the state to the already intensive profiling that can be done on individuals in the market. While state data published as open government data would be aggregate or de-identified, it may still include sensitive information on several other axes, which, when combined with commercial knowledge about an individual, could add to prejudicial data profiling.

The counter-argument to this risk is that individuals belonging to aggregate groups with positive characteristics, in the context of accessing services, should be offered the benefit of this profiling, often characterized by access, or less expensive access, to a set of products or services. This argument is weakened when there are factors in play that are beyond an individual's control, such as health conditions. Besides, factors *within* their control should still not be taken as grounds for prejudicial access to goods and services. In any case, the trade-off here to consider is whether opening and publishing any type of data that may increase prejudicial service delivery is worth any possible upside to a set of individuals. The state will eventually have to bear the costs to support individuals that are rejected from market access to goods and services. Insurers, for example, are a type of financial government and any ways that open data policy might further entrench their power ought to be considered, based on the fact that the state is ultimately

the one left to support anyone that does not make measure as a customer (Attali 2009).

As recently announced plans in France and Ontario illustrate (Villani 2018),<sup>3</sup> open health-care data may be released to support scientific research and AI. Just as the upside may be to improve capacity to identify and pre-empt disease, there is also an increased risk of prejudicial application of the data. The upside on the medical research front may simply not be worth the potential negative effects on the public, in particular if those that are marginalized face further discrimination in accessing health care or services or are underrepresented in the data. In cases related to scientific research, aggregate or de-identified open data related to health may be best served by cordoning it off to a shared data site where it can only be used for research and medical purposes.<sup>4</sup> The challenge here, however, is that it becomes impossible to follow the trail and use of data, in particular when it is in a commercial context with no reason or regulation to share how data is used in product and service delivery. The technology sector is not currently held to the same ethical data-use protocols that are used in scientific research or the public service. Here we return to limitations within privacy and ethics, in how the frame and lens applied in law and policy is set to the individual level and is too narrow to capture some of these societal and collective concerns regarding the aggregate use of data (Vedder 1999).

The second issue with this set of impacts is that it broadens a power asymmetry. According to Mireille Hildebrandt (2008), “The most salient difference between organic and machine profiling may be the fact that as a citizen, consumer or employee we find ourselves in the position of being profiled, without access to the knowledge that is used to categorise and deal with us. This seems to impair our personal freedom, because we cannot adequately anticipate the actions of those that know about us what we may not know about ourselves.”

To exist in the face of discriminatory behaviour is one thing, but to not know about it in order to address it is on another level completely. Take the history of personal credit as an example. The institutions that manage personal credit are heavily regulated (Scassa 2015), and they are required to

---

3 See <https://projectspark.ca/about-spark/#>.

4 See <https://www.saluscoop.org/>.

make information about an individual's credit history, credit score and rationale for ratings accessible upon request. Yet, based on a Canadian court ruling in 2003, Canada's commercial privacy act, PIPEDA, found that an individual was unable to gain access to a bank's internal data process, an algorithm that was used to make a decision about granting them access to a credit product. The ruling found that sharing this information about how the algorithm worked could have a negative commercial impact on the bank. This is a less-than-encouraging precedent ruling in Canada for algorithmic accountability (*ibid.*).

Consider this issue in the modern context, where the range of additional measures used to profile customers within private sector contexts has multiplied immensely. We can be profiled across many axes. In any modern-day equivalent, how is one to address or update or "fix" one's data profile if one has completely lost track of what that profile is and may not be able to gain access to that profile in the first place? More interestingly, what is an individual to do when that profile is not fixed nor of one version — if the profile is constantly being revised and remixed as different sets of data are brought together.

There is no doubt a set of heavy trade-offs to consider. As social scientists have long argued, there is a need to access better and higher-quality data to assist in the development of social service delivery, and in other cases. There is a large set of stakeholders that have sought open data for good and ethical reasons and hindering their work would be counterproductive. However, there are still other options on the open data spectrum to consider. In cases where there is a scientific and non-commercial research interest, research data centres and other approaches may hold the key to a middle group — opening access enough to support improvements and science, but not to the totally unmitigated, impossible-to-track level of fully open data.

The final case to consider with open data is an urban planning example, which presents itself in cases related to traffic and pattern optimization. Google's Waze product allows drivers to access a tool built on a shared open data set to predict an optimum route. Recent studies show these types of apps may be making traffic worse (Madrigal 2018). These apps are framed as beneficial due not only to their utility but have also been partially "sold" via their commitment to providing open data, in a

sharing and partnership model with government (Rider 2017), and with a side feature of emergency messaging that may further entrench the app within government operations. Waze's business model is location-based targeted advertising. On its website, the first line in its "About Us" section states: "Waze is all about contributing to the 'common good' out there on the road."<sup>5</sup>

Almost 30 years ago, various researchers began to challenge the assumption that providing better information to drivers would resolve congestion (Mahmassani et al. 1991; Arnott, de Palma and Lindsey 1991). The benefits of this open data do not accrue to the public equally. First, they may provide value to those who can use the data due to digital infrastructure and access. Then, as suggested by Hani S. Mahmassani et al. (1991), there may be cases where the entire model is counterproductive — where the idea of data collection and its use does not stand up to scientific rigour. This draws to mind the data assemblage model, where data is collected and organized into a product and sold in a corporate frame without any scientific rigour to test the hypothesis that all the additional data in the hands of individuals actually improves system design (Kitchin and Lauriault 2014). This may be a case where there is unequal enough power at the moment to keep all systems functioning, but if all users were empowered with the data and product, it could lead to persistent system breakdown. This introduces the challenge of having to try and "walk back" an innovation — an idea that users have become wed to. How is one to take back a resident-generated data system that has taken hold and can operate entirely between the user/resident and the corporation? The possible mechanisms are regulatory. A final consideration in this model may be the additional negative impacts on those who are not participating in the data model but suffer service quality reduction on public transit due to the introduction of technology when its impacts were not understood.

The summary consideration here is that within our liberal democracy there are many models that depend on averaging and shared risk. A degree of imprecision is what enables many of our systems to work. Where there is a commercial interest, collective human interest will never be protected at the cost of profit reduction without regulation. For every mention of an individualistic

---

5 See [www.waze.com/about](http://www.waze.com/about).

right or protection, a collective mechanism must be raised for consideration. How do we consider aggregate needs inherent to the shared risk model that liberal democratic markets are built on? Liberal democracy relies on a collective notion of humanity, public service and good governance. For every mention of privacy by design and concerns for individual and personally identifiable information, there must be made new mention of data minimization, mass consent, aggregate privacy and privacy as a public good. For every mention of efficiency, there must be made mention of trade-offs related to universal protections. For every effort to privatize a service using data or technology, there must be a discussion about the impacts of the blurring lines between service delivery by the state versus the market.

We are moving along a continuum where the collective notion of how to participate in a system as a community is decreasing. Micro-targeting and individual services have been a boon to many in the commercial space and they continue to push this slide toward individually focused behaviours and norms. Now we are facing new technology issues that are triggering an exploration of the downsides associated with these types of individually focused systems. Current circumstances may create a reminder of the need to give more than one receives, in some cases, in order to support the full range of human existence in our society. And, possibly harder, to bypass some of the opportunities and technologies being made available as a necessary step to protect ourselves as a collective.

---

## Policy Guidance

These are a brief set of unformed but high-level suggestions to consider in a policy development context when considering human behavioural data and open data programs moving forward:

- Create a process to assess and consider data that is tied to humans prior to open data release. This work may include a categorization exercise to determine which human data is most sensitive and help inform protocols in a related manner.
- Mandate government ownership for data collected in public spaces and

build the requirement into procurement rules. This ownership can then be used to democratically guide usage through licensing, including open data.

- Define types of data that should not be collected at all and create guidelines to drive this decision. This may include repealing allowances for current practices. The only secure human data is data that has not been captured. A public consultation should be held to determine not only what data to capture and why, but also which data not to capture, and why not. This would include potentially rolling back some of the existing programs and allowances within policies.
- Suggest updates to open-by-default policy/ open data charter to reflect these possible issues with human behavioural data and/ or convene discussions through the open government partnership to explore them further within the open government community.
- Convene a national discussion on data justice and data rights.
- Fund community infrastructure programs to support the hardware, software, governance and community capacity requirements for communities to oversee and manage their own data collection.

---

## Conclusion

Given the many unexpected, unknown and unintentional outcomes regarding the use of data, the technologically deterministic approach inherent to the belief in open data, like technology, as inherently good or an inevitable state of affairs needs a rethink. Considering the particular implications of the use of human behavioural data to further commoditize human behaviour may be helpful in this exercise. There are cultural influences at play in our current data discourse: habits and behavioural evolutions or norms that demand their own distinct analysis, including a rethink of ideas related to consent and social licence. The impact of the world moving away from modelling and planning and into actuals is profound. In some ways, there is both safety and

imprecision in inefficient and complex models. There is an important tension and trade-offs discussion to be had about whether the state's importance as an institution of stability may need to be heavily re-exerted, and to create a space for an exploration of social values. This paper is an effort to frame and inform these discussions. Some trade-offs and loss of efficiency in service delivery may be required to protect the notion of free will and places that exist without surveillance and/or plausible surveillance. Either we work to find democratically informed and intentional approaches to thrive using data and technology or we move aggressively into the collapse of the remaining state institutions and adopt a boldly techno-deterministic and libertarian agenda.

## Author's Note

This paper is the product of talking, thinking and working with a number of great colleagues and collaborators. A special thank you to Morgan MacInnes at CIGI for work on the early drafts. Many thanks also to: Nabeel Ahmed, Nasma Ahmed, Andrew Do, Amira Elghawaby, Melissa Goldstein, Meghan Hellstern, Kailee Hilt, Jury Konga, Tracey Lauriault, Alex Lougheed, Bryan McHale, James McKinney, Andrea Migone, David Murakami Wood, Saadia Muzaffar, Ushnish Sengupta, Pamela Robinson, Teresa Scassa, Renee Sieber, Ashleigh Weeden and Beth Wilson. I am very thankful for their willingness to think with me on this topic, and anything that is wrong with this paper rests squarely and solely on me. Finally, the biggest thank you of all goes to Mark, Max and Lana — my loves and my home.

---

## Works Cited

- Argast, Andi and Lydia Zvyagintseva. 2016. "Data Literacy projects in Canada: field notes from the Open Data Institute, Toronto node." *The Journal of Community Informatics* 12 (3): 223-31. <http://ci-journal.org/index.php/ciej/article/viewFile/1302/1227>.
- Arnott, Richard, Andre de Palma and Robin Lindsey. 1991. "Does providing information to drivers reduce traffic congestion?" *Transportation Research Part A: General* 25 (5): 309-18. [https://doi.org/10.1016/0191-2607\(91\)90146-H](https://doi.org/10.1016/0191-2607(91)90146-H).
- Attali, Jacques. 2009. *A Brief History of the Future: A Brave and Controversial Look at the Twenty-first Century*. Crows Nest, New South Wales, Australia: Allen & Unwin.
- Bailey, Stuart G. M. and Nadia Caidi. 2005. "How much is too little? Privacy and smart cards in Hong Kong and Ontario." *Journal of Information Science* (October). DOI: 10.1177/0165551505055400.
- Bates, Jo. 2012. "This is what modern deregulation looks like: Co-optation and contestation in the shaping of the UK's Open Government Data Initiative." *The Journal of Community Informatics* 8 (2). <http://ci-journal.org/index.php/ciej/article/view/845/916>.
- Berger, Andrew. 2009. "Brandeis and the History of Transparency." *Sunlight Foundation* (blog), May 26. <https://sunlightfoundation.com/2009/05/26/brandeis-and-the-history-of-transparency/>.
- Brandusescu, Ana, Renée E. Sieber and Sylvie Jochems. 2015. "Confronting the hype: The use of crisis mapping for community development." *Convergence: The International Journal of Research into New Media Technologies* 22 (6): 616-32.
- Caidi, Nadia and Anthony Ross. 2005. "Information Rights and National Security." *Government Information Quarterly* 22: 663-84. doi:10.1016/j.giq.2005.11.003.
- Chomsky, Noam and Michael Foucault. 2006. *The Chomsky-Foucault Debate on Human Nature*. New York, NY: The New Press.



- Citron, Danielle Keats. 2007. "Technological Due Process." University of Maryland Legal Studies Research Paper No. 2007-26. *Washington University Law Review* 85: 1249-1313. <https://ssrn.com/abstract=1012360>.
- Ciuriak, Dan. 2018. "The Economics of Data: Implications for the Data-driven Economy." In *Data Governance in the Digital Age*, 12-19. Waterloo, ON: CIGI. [www.cigionline.org/publications/data-governance-digital-age](http://www.cigionline.org/publications/data-governance-digital-age).
- David, Paul A. 2008. "The Historical Origins of 'Open Science': An Essay on Patronage, Reputation and Common Agency Contracting in the Scientific Revolution." *Capitalism and Society* 3 (2), article 5. <https://ssrn.com/abstract=2209188>.
- Durkheim, Émile. 1957. *Professional Ethics and Civic Morals*. London, UK: Routledge.
- Edwards, Lilian. 2016. "Privacy, Security and Data Protection in Smart Cities: A Critical EU Law Perspective." *European Data Protection Law Review*, January 5. <http://dx.doi.org/10.2139/ssrn.2711290>.
- Esguerra, Richard. 2009. "Google CEO Eric Schmidt Dismisses the Importance of Privacy." *Electronic Frontier Foundation* (blog), December 10. [www.eff.org/deeplinks/2009/12/google-ceo-eric-schmidt-dismisses-privacy](http://www.eff.org/deeplinks/2009/12/google-ceo-eric-schmidt-dismisses-privacy).
- Eubanks, Virginia. 2017. *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*. New York, NY: St. Martin's Press.
- European Commission. 2017. "Staff Working Document on the free flow of data and emerging issues of the European data economy." January 10. <https://ec.europa.eu/digital-single-market/en/news/staff-working-document-free-flow-data-and-emerging-issues-european-data-economy>.
- Fairfield, Joshua A. T. and Christoph Engel. 2015. "Privacy as a Public Good." *Duke Law Journal* 65: 385-457. <https://scholarship.law.duke.edu/dlj/vol65/iss3/1>.
- Foroohar, Rana. 2018. "The Dangers of Digital Democracy." *The Financial Times*, March 6. [www.ft.com/content/356140e0-02ac-11e8-9650-9c0ad2d7c5b5](http://www.ft.com/content/356140e0-02ac-11e8-9650-9c0ad2d7c5b5).
- Frizell, Sam. 2014. "What Is Uber Really Doing With Your Data?" *Time Magazine*, November 19. <http://time.com/3595025/uber-data/>.
- Gill, Lex, Dennis Redeker and Urs Gasser. 2015. "Towards Digital Constitutionalism? Mapping Attempts to Craft an Internet Bill of Rights." Berkman Center Research Publication No. 2015-15. <http://dx.doi.org/10.2139/ssrn.2687120>.
- Goldsmith, Stephen and Neil Kleiman. 2017. *A New City O/S: The Power of Open, Collaborative, and Distributed Governance*. Washington, DC: Brookings Institution Press.
- Gurstein, Michael. 2011. "Open data: Empowering the empowered or effective data use for everyone?" *First Monday* 16 (2). <http://ojphi.org/ojs/index.php/fm/article/view/3316/2764>.
- Hildebrandt, Mireille. 2008. "Defining Profiling: A New Type of Knowledge?" In *Profiling the European Citizen*, edited by M. Hildebrandt and S. Gutwirth, 17-45. Dordrecht, Netherlands: Springer.
- Howard, Alex. 2016. "10 takeaways from the 2016 International Open Data Conference." *Sunlight Foundation* (blog), October 7. <https://sunlightfoundation.com/2016/10/07/10-takeaways-from-the-2016-international-open-data-conference/>.
- Howlett, Michael and Andrea Migone. 2014. "Making the invisible public service visible? Exploring data on the supply of policy and management consultancies in Canada." *Canadian Public Administration* 57 (2): 183-216. <https://onlinelibrary.wiley.com/toc/17547121/57/2>.
- Ireton, Julie. 2018. "Cost of Phoenix federal payroll debacle surpasses \$1B." *CBC News*, March 26. [www.cbc.ca/news/canada/ottawa/phoenix-cost-more-than-one-billion-dollars-1.4594115](http://www.cbc.ca/news/canada/ottawa/phoenix-cost-more-than-one-billion-dollars-1.4594115).
- Janssen, Katleen. 2012. "Open Government Data: Right to Information 2.0 or its Rollback Version?" ICRI Research Paper No. 8. September 1. <https://ssrn.com/abstract=2152566>.
- Janssen, Marijn and George Kuk. 2016. "The challenges and limits of big data algorithms in technocratic governance." *Government Information Quarterly* 33 (3): 371-77. doi:10.1016/j.giq.2016.08.011.

- Johnson, Jeffrey Alan. 2014. "From open data to information justice." *Ethics and Information Technology* 16 (4): 263-74. doi:10.1007/s10676-014-9351-8.
- Johnson, Peter, Renee Sieber, Teresa Scassa, Monica Stephens and Pamela J. Robinson. 2017. "The Cost(s) of Geospatial Open Data." *Transactions in GIS* 21 (3): 434-45. DOI: 10.1111/tgis.12283.
- Kahn, Jeremy and Alex Morales. 2018. "U.K. Unveils \$1.4 Billion Drive Into Artificial Intelligence." Bloomberg, April 25. [www.bloomberg.com/news/articles/2018-04-25/u-k-announces-1-4-billion-drive-into-artificial-intelligence](http://www.bloomberg.com/news/articles/2018-04-25/u-k-announces-1-4-billion-drive-into-artificial-intelligence).
- Kissinger, Henry. A. 2018. "How the Enlightenment Ends." *The Atlantic*, June. [www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/](http://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/).
- Kitchin, Rob. 2013. "The real-time city? Big data and smart urbanism." *GeoJournal* 79 (1): 1-14. DOI 10.1007/s10708-013-9516-8.
- Kitchin, Rob and Tracey Lauriault. 2014. "Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work." The Programmable City Working Paper 2, July 30; pre-print version of chapter to be published in *Thinking Big Data in Geography: New Regimes, New Research*, edited by J. Thatcher, A. Shears and J. Eckert, 3-20. University of Nebraska Press. <https://ssrn.com/abstract=2474112>.
- Kitchin, Rob, Tracey Lauriault and Gavin McArdle. 2015. "Smart cities and the politics of urban data." In *Smart Urbanism: Utopian Vision or False Dawn?*, edited by Simon Marvin, Andrés Luque-Ayala and Colin McFarlane, 16-33. London, UK: Routledge.
- Kline, Ronald R. 2015. "Technological Determinism." In *International Encyclopedia of the Social & Behavioral Sciences*. 2nd ed. Vol. 24. <http://dx.doi.org/10.1016/B978-0-08-097086-8.85034-5>.
- Kwecka, Zbigniew, William Buchanan, Burkhard Schafer and Judith Rauhofer. 2014. "'I am Spartacus': privacy enhancing technologies, collaborative obfuscation and privacy as a public good." *Artificial Intelligence and Law* 22 (2): 113-39. DOI 10.1007/s10506-014-9155-5.
- Landry, Jean-Noé, Kara Webber, Bianca Wylie and Pamela Robinson. 2017. "How can we improve urban resilience with open data?" Open Data for Development. [www.ccmdesign.ca/files/od4d-resilient-cities.pdf](http://www.ccmdesign.ca/files/od4d-resilient-cities.pdf).
- Lauriault, Tracey, Rachel Bloom and Jean-Noé Landry. 2018. "Open Smart Cities Guide V1.0." Natural Resources Canada, GeoConnections Program. [www.opennorth.ca/open-smart-cities-guide](http://www.opennorth.ca/open-smart-cities-guide).
- Lessig, Lawrence. 2000. "Code Is Law: On Liberty in Cyberspace." *Harvard Magazine*, January 1. <https://harvardmagazine.com/2000/01/code-is-law-html>.
- Longo, Justin. 2011. "#Opendata: Digital-Era Governance Thoroughbred or New Public Management Trojan Horse?" *Public Policy & Governance Review* 2 (2): 38. <https://ssrn.com/abstract=1856120>.
- Longo, Justin, Evan Kuras, Holly Smith, David M. Hondula and Erik Johnston. 2017. "Technology Use, Exposure to Natural Hazards, and Being Digitally Invisible: Implications for Policy Analytics." *Policy & Internet* 9 (1): 76-108. [https://jilphd.files.wordpress.com/2017/02/longo\\_et\\_al-2017-policy\\_internet.pdf](https://jilphd.files.wordpress.com/2017/02/longo_et_al-2017-policy_internet.pdf).
- Lyon, David. 2017. "Surveillance Culture: Engagement, Exposure, and Ethics in Digital Modernity." *International Journal of Communication* 11: 824-42.
- Madrigal, Alexis C. 2018. "The Perfect Selfishness of Mapping Apps." *The Atlantic*, March 15. [www.theatlantic.com/technology/archive/2018/03/mapping-apps-and-the-price-of-anarchy/555551/](http://www.theatlantic.com/technology/archive/2018/03/mapping-apps-and-the-price-of-anarchy/555551/).
- Mahmassani, Hani S., Srinivas Peeta, Gang-Len Chang and Thanavat Junchaya. 1991. "A Review of Dynamic Assignment and Traffic Simulation Models for ADIS/ATMS Applications." Center for Transportation Research at the University of Austin. March. <https://library.ctr.utexas.edu/digitized/texasarchive/phase2/mahmassani-1991.pdf>.
- Maley, Terry. 2004. "Max Weber and the Iron Cage of Technology." *Bulletin of Science, Technology & Society* 24 (1): 69-86. <http://journals.sagepub.com/doi/10.1177/0270467604263181>.

- Manyika, James, Michael Chui, Peter Groves, Diana Farrell, Steve Van Kuiken and Elizabeth Almasi Doshi. 2013. *Open Data: Unlocking Innovation and Performance with Liquid Information*. October. McKinsey Global Institute. [www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information](http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information).
- Marshall, Ellie. 2018. "Legitimate Invasions: What Ontario can learn from the history of the Consumer Reporting Act" (unpublished paper).
- Mattern, Shannon. 2017. "A City Is Not a Computer." *Places Journal* (February). <https://placesjournal.org/article/a-city-is-not-a-computer/>.
- . 2018. "Databodies in Codespace." *Places Journal* (April). <https://placesjournal.org/article/databodies-in-codespace/>.
- Metz, Julia. 2015. *The European Commission, Expert Groups, and the Policy Process: Demystifying Technocratic Governance*. New York, NY: Palgrave Macmillan.
- Meyer, Michelle. 2014. "Everything You Need to Know about Facebook's Controversial Emotion Experiment." *Wired*, June 30. [www.wired.com/2014/06/everything-you-need-to-know-about-facebooks-manipulative-experiment/](http://www.wired.com/2014/06/everything-you-need-to-know-about-facebooks-manipulative-experiment/).
- Murakami Wood, David. 2013. "What is global surveillance? Towards a relational political economy of the global surveillant assemblage." *Geoforum* (49): 317-26.
- . 2015. "Smart City, Surveillance City." *Society for Computers and Law*. [www.scl.org/articles/3405-smart-city-surveillance-city](http://www.scl.org/articles/3405-smart-city-surveillance-city).
- Murakami Wood, David and Jennie Day. 2018. "Dipping into the Data Flow: Surveillance by Proxy in Smart Urban Environments" (unpublished paper).
- Murakami Wood, David and Kirstie Ball. 2013. "Brandscapes of control? Surveillance, marketing and the co-construction of subjectivity and space in neo-liberal capitalism." *Marketing Theory* 1-21.
- Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York, NY: NYU Press.
- Obar, Jonathan A. 2015. "Big Data and The Phantom Public: Walter Lippmann and the fallacy of data privacy self-management." *Big Data & Society* 2 (2): 1-16.
- O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY: Crown Publishers.
- Open Data Institute. 2017. "What is 'open data' and why should we care?" *Open Data Institute* (blog), February 3. <https://theodi.org/article/what-is-open-data-and-why-should-we-care/>.
- O'Reilly, Tim. 2009. "Gov 2.0: The Promise Of Innovation." *Forbes*, August 10. [www.forbes.com/2009/08/10/government-internet-software-technology-breakthroughs-oreilly.html#d287aec3b7b4](http://www.forbes.com/2009/08/10/government-internet-software-technology-breakthroughs-oreilly.html#d287aec3b7b4).
- Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press.
- . 2018. "Tech Platforms and the Knowledge Problem." *American Affairs Journal* II (2). <https://americanaffairsjournal.org/2018/05/tech-platforms-and-the-knowledge-problem/>.
- Popper, Karl. 1945. *The Open Society and Its Enemies*. London, UK: Routledge.
- Ranchordas, Sofia and Abram Klop. Forthcoming 2018. "Data-Driven Regulation and Governance in Smart Cities." In *Research Handbook in Data Science and Law*, edited by Vanessa Mak, Eric T. T. Tai and Anna Berlee. Cheltenham, UK: Edward Elgar.
- Rider, David. 2017. "Toronto and Waze app agree to trade traffic data." *The Toronto Star*, November 20. [www.thestar.com/news/city\\_hall/2017/11/20/toronto-and-waze-app-agree-to-trade-traffic-data.html](http://www.thestar.com/news/city_hall/2017/11/20/toronto-and-waze-app-agree-to-trade-traffic-data.html).
- Robinson, Pamela. 2018. "Design Interventions for a Digital Venture." *Spacing* (spring).
- Robinson, Pamela and Dina Graser. 2016. "A Recipe for Fiscal Trust." *IMFG Perspectives*, no. 13. [https://munkschool.utoronto.ca/imfg/uploads/350/imfgperspectives\\_no13\\_fiscaltrust\\_graser\\_robinson\\_july\\_21\\_2016.pdf](https://munkschool.utoronto.ca/imfg/uploads/350/imfgperspectives_no13_fiscaltrust_graser_robinson_july_21_2016.pdf).

- Robinson, Pamela and Lisa Ward-Mather. 2017. "Open Data Community Maturity: Libraries as Civic Infomediaries." *Journal of the Urban and Regional Information Science Association Journal* 28: 31-39. [www.urisa.org/clientuploads/directory/Documents/Journal/Vol28\\_final.pdf](http://www.urisa.org/clientuploads/directory/Documents/Journal/Vol28_final.pdf).
- Ruijter, Erna, Stephan Grimmelikhuijsen and Albert Meijer. 2017. "Open data for democracy: Developing a theoretical framework for open data use." *Government Information Quarterly* 34 (1): 45-52. [www.sciencedirect.com/science/article/pii/S0740624X17300011](http://www.sciencedirect.com/science/article/pii/S0740624X17300011).
- Sánchez, David and Alexandre Viejo. 2017. "Personalized privacy in open data sharing scenarios." *Online Information Review* 41 (3): 298-310. <https://doi.org/10.1108/OIR-01-2016-0011>.
- Scannell, R. Joshua. 2016. "Broken Windows, Broken Code." *Reallifemag.com*, August 29. <http://reallifemag.com/broken-windows-broken-code/>.
- Scassa, Tersea. 2014. "Privacy and Open Government." *Future Internet* 6 (2): 397-413. doi:10.3390/fi6020397.
- . 2015. "Back to the Future I: What Past Privacy Findings Tell us About the Future of Big Data and Privacy." March 2. [www.teresascassa.ca/index.php?option=com\\_k2&view=item&id=180:back-to-the-future-i-what-past-privacy-findings-tell-us-about-the-future-of-big-data-and-privacy&Itemid=80](http://www.teresascassa.ca/index.php?option=com_k2&view=item&id=180:back-to-the-future-i-what-past-privacy-findings-tell-us-about-the-future-of-big-data-and-privacy&Itemid=80).
- . 2018. "Considerations for Canada's National Data Strategy." In *Data Governance in the Digital Age*. Waterloo, ON: CIGI. [www.cigionline.org/publications/data-governance-digital-age](http://www.cigionline.org/publications/data-governance-digital-age).
- . Forthcoming 2018. *Data Ownership*. CIGI Paper No. 187. Waterloo, ON: CIGI.
- Sieber, Renee E. and Peter Johnson. 2015. "Civic open data at a crossroads: Dominant models and current challenges." *Government Information Quarterly* 32 (3): 308-15. [www.sciencedirect.com/science/article/pii/S0740624X15000611](http://www.sciencedirect.com/science/article/pii/S0740624X15000611).
- Solon, Olivia. 2018. "Cambridge Analytica whistleblower says Bannon wanted to suppress voters." *The Guardian*, May 16. [www.theguardian.com/uk-news/2018/may/16/steve-bannon-cambridge-analytica-whistleblower-suppress-voters-testimony](http://www.theguardian.com/uk-news/2018/may/16/steve-bannon-cambridge-analytica-whistleblower-suppress-voters-testimony).
- Sunstein, Cass R. and Richard H. Thaler. 2005. "Libertarian Paternalism is Not an Oxymoron." *University of Chicago Law Review* 99 (3). <https://ssrn.com/abstract=615562>.
- Taylor, Linnet. 2017. "What is data justice? The case for connecting digital rights and freedoms globally." *Big Data & Society*, November 1. <https://doi.org/10.1177/2053951717736335>.
- Tennison, Jeni. 2018. "Doesn't open data make data monopolies more powerful?" January 14. [www.jenitennison.com/2018/01/14/data-monopolies.html](http://www.jenitennison.com/2018/01/14/data-monopolies.html).
- Thereaux, Olivier. 2017. "Using artificial intelligence and open data for innovation and accountability." *Open Data Institute (blog)*, December 20. <https://theodi.org/article/using-artificial-intelligence-and-open-data-for-innovation-and-accountability/>.
- Thierer, Adam. 2014. *Permissionless Innovation: The Continuing Case for Comprehensive Technological Freedom*. Arlington, VA: Mercatus Center, George Mason University.
- Thompson, Nicholas. 2018. "Emmanuel Macron Talks to WIRED About France's AI Strategy." WIRED, March 31. [www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy/](http://www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy/).
- Tufekci, Zeynep. 2014. "Engineering the public: Big data, surveillance and computational politics." *First Monday* 19 (7). <http://dx.doi.org/10.5210/fm.v19i7.4901>.
- Varian, Hal. 2014. "Beyond Big Data." *Business Economics* 49 (1): 27-31.
- Vedder, Anton. 1999. "KDD: The challenge to individualism." *Ethics and Information Technology* (1): 275-281. <https://doi.org/10.1023/A:1010016102284>.

- Villani, Cédric. 2018. *For a Meaningful Artificial Intelligence: Towards a French and European Strategy*. [www.aiforhumanity.fr/pdfs/MissionVillani\\_Report\\_ENG-VF.pdf](http://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf).
- Waldman, Peter, Lizette Chapman and Jordan Robertson. 2018. "Palantir Knows Everything About You." Bloomberg, April 19. [www.bloomberg.com/features/2018-palantir-peter-thiel/](http://www.bloomberg.com/features/2018-palantir-peter-thiel/).
- Weerakkody, Vishanth, Zahir Irani, Kawal Kapoor, Uthayasankar Sivarajah and Yogesh K. Dwivedi. 2017. "Open data and its usability: an empirical view from the Citizen's perspective." *Information Systems Frontier* 19 (2): 285-300. <https://doi.org/10.1007/s10796-016-9679-1>.
- Wolpert, Ann J. 2013. "For the Sake of Inquiry and Knowledge — The Inevitability of Open Access." *The New England Journal of Medicine* 368: 785-87. DOI: 10.1056/NEJMp1211410.
- Wylie, Bianca. 2017. "Countering the Digital Consensus — Technology, Elites, and Elections." *Code for America* (blog), June 27. <https://medium.com/code-for-america/countering-the-digital-consensus-technology-elites-and-elections-b866ee5cde51>.
- . 2018a. "Governance Vacuums and How Code Is Becoming Law." In *Data Governance in the Digital Age*. Waterloo, ON: CIGI. [www.cigionline.org/publications/data-governance-digital-age](http://www.cigionline.org/publications/data-governance-digital-age).
- . 2018b. "Searching for the Smart City's Democratic Future." [www.cigionline.org/articles/searching-smart-citys-democratic-future](http://www.cigionline.org/articles/searching-smart-citys-democratic-future).
- Yeung, Karen. 2016. "'Hypernudge': Big Data as a Mode of Regulation by Design." *Information, Communication and Society* 20 (1): 1-19. <https://ssrn.com/abstract=2807574>.
- Zeleti, Fatemeh Ahmadi, Adegboyega Ojo and Edward Curry. 2016. "Exploring the economic value of open government data." *Government Information Quarterly* 33 (3): 535-51. <https://doi.org/10.1016/j.giq.2016.01.008>.
- Zuboff, Shoshana. 2014. Keynote: Reality is the Next Big Thing — Elevate Festival. YouTube. [www.youtube.com/watch?v=0QwPHinDdOc](http://www.youtube.com/watch?v=0QwPHinDdOc).
- . 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." *Journal of Information Technology* 30 (1): 75-89. doi:10.1057/jit.2015.5.
- Zuiderwijk, Anneke, Rhythmia Shinde and Marijn Janssen. 2018. "Investigating the attainment of open government data objectives: Is there a mismatch between objectives and results?" *International Review of Administrative Sciences*, February 8. <https://doi.org/10.1177/0020852317739115>.

---

## About CIGI

We are the Centre for International Governance Innovation: an independent, non-partisan think tank with an objective and uniquely global perspective. Our research, opinions and public voice make a difference in today's world by bringing clarity and innovative thinking to global policy making. By working across disciplines and in partnership with the best peers and experts, we are the benchmark for influential research and trusted analysis.

Our research programs focus on governance of the global economy, global security and politics, and international law in collaboration with a range of strategic partners and support from the Government of Canada, the Government of Ontario, as well as founder Jim Balsillie.

---

## À propos du CIGI

Au Centre pour l'innovation dans la gouvernance internationale (CIGI), nous formons un groupe de réflexion indépendant et non partisan doté d'un point de vue objectif et unique de portée mondiale. Nos recherches, nos avis et nos interventions publiques ont des effets réels sur le monde d'aujourd'hui car ils apportent de la clarté et une réflexion novatrice pour l'élaboration des politiques à l'échelle internationale. En raison des travaux accomplis en collaboration et en partenariat avec des pairs et des spécialistes interdisciplinaires des plus compétents, nous sommes devenus une référence grâce à l'influence de nos recherches et à la fiabilité de nos analyses.

Nos programmes de recherche ont trait à la gouvernance dans les domaines suivants : l'économie mondiale, la sécurité et les politiques mondiales, et le droit international, et nous les exécutons avec la collaboration de nombreux partenaires stratégiques et le soutien des gouvernements du Canada et de l'Ontario ainsi que du fondateur du CIGI, Jim Balsillie.



---

**Centre for International  
Governance Innovation**

67 Erb Street West  
Waterloo, ON, Canada N2L 6C2  
[www.cigionline.org](http://www.cigionline.org)

 @cigionline

