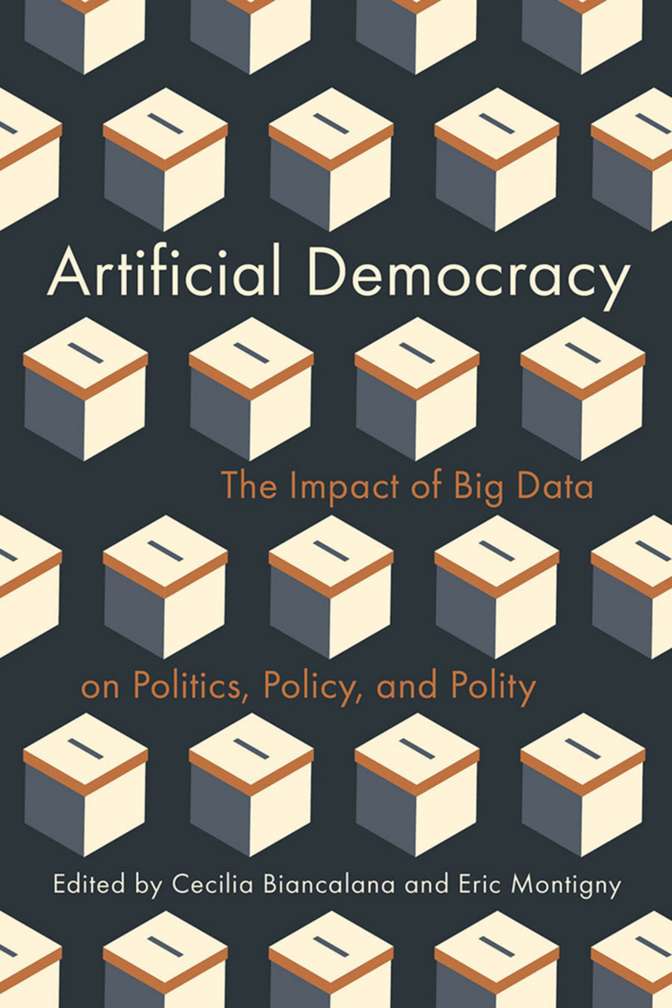

Artificial Democracy

The Impact of Big Data on Politics, Policy, and Polity

Edited by Cecilia Biancalana and Eric Montigny

Artificial Democracy

This page intentionally left blank

Artificial Democracy

The Impact of Big Data on Politics, Policy, and Polity

Edited by Cecilia Biancalana
and Eric Montigny

UBCPress · Vancouver

UBC Press is a Benetech Global Certified Accessible™ publisher. The epub version of this book meets stringent accessibility standards, ensuring it is available to people with diverse needs.

Library and Archives Canada Cataloguing in Publication

Title: Artificial democracy : the impact of big data on politics, policy, and polity / edited by Cecilia Biancalana and Eric Montigny.

Names: Biancalana, Cecilia, editor. | Montigny, Éric, editor

Description: Includes bibliographical references and index.

Identifiers: Canadiana (print) 20250163179 | Canadiana (ebook) 20250163195 | ISBN 9780774871013 (hardcover) | ISBN 9780774871037 (PDF) | ISBN 9780774871044 (EPUB)

Subjects: LCSH: Democracy. | LCSH: Big data—Political aspects. | LCSH: Artificial intelligence—Political aspects.

Classification: LCC JC423 .A78 2025 | DDC 321.8—dc23

Canada Council for the Arts
Conseil des arts du Canada

Canada

British Columbia Arts Council

British Columbia

UBC Press gratefully acknowledges the financial support for our publishing program of the Government of Canada, the Canada Council for the Arts, and the British Columbia Arts Council.

This book has been published with the help of a grant from the Canadian Federation for the Humanities and Social Sciences, through the Scholarly Book Awards, using funds provided by the Social Sciences and Humanities Research Council of Canada.

UBC Press is situated on the traditional, ancestral, and unceded territory of the xʷməθkʷəy̓əm (Musqueam) people. This land has always been a place of learning for the xʷməθkʷəy̓əm, who have passed on their culture, history, and traditions for millennia, from one generation to the next.

UBC Press
The University of British Columbia
www.ubcpress.ca

Contents

Part 1: Polity: A Regulatory Framework for a Democratic Use of Data and AI

Part 2: Politics: Use of Data, Profiling, and Personalizing in Electioneering

Part 3: Policy: Surveillance and Data Protection during the Coronavirus Pandemic

Artificial Democracy

This page intentionally left blank

Introduction

Towards an Artificial Democracy?

Eric Montigny and Cecilia Biancalana

Are Democracies Being Transformed?

These are hard times for representative democracy. According to the International Institute for Democracy and Electoral Assistance (IDEA, 2022), nearly half of the world's democracies are now in decline – 48 out of 104. At the same time, the Sweden-based organization points out that authoritarian regimes are even more repressive than before (IDEA, 2022). The 2022 World Values Survey data (which includes seventy-seven jurisdictions) also reveals a decline in the attractiveness of democracy for citizens. Now, 52 percent say they prefer to be governed by a strong leader who does not have to bother with Parliament or elections, up from 38 percent in 2009 (DemDigest, 2022).

The symptoms of this alleged "democratic crisis," a term popular since the 1970s in Western democracies, would also encompass declining voter turnout, decreasing party memberships, a rise in electoral volatility indicating that traditional cleavages are thawing, as well as the emergence and success of new parties: anti-establishment, antisystem, or populist. These phenomena collectively reflect a growing disillusionment or discontent among citizens with established political processes and institutions, leading to shifts in voting patterns and the proliferation of alternative political movements.

Yet, in the aftermath of the fall of the Berlin Wall in 1989, alternatives to democracy seemed doomed to die. A wave of democratization was sweeping through Eastern Europe and sub-Saharan Africa. It had also happened in South America. Nothing seemed to be able to stop the march towards democracy. In 2012, a new wave of democratization seemed to carry the Middle East and North Africa. This time, what was called the Arab Spring was said to be fuelled by social networks. These new digital tools, it was believed, would prove to be powerful for citizen mobilization. Following the spread of the internet, dictatorships were threatened and destined to disappear.

But what role does technology really play in either bolstering or undermining democracies? Technologies are not just tools; they shape how we perceive the world, thereby exerting a significant social influence. However, it is crucial to steer clear of two pitfalls: "technological determinism," the belief that technology alone drives change, and "technological solutionism," the mistaken notion that technology is a panacea for all political problems. Instead, we must recognize that technology and media serve as "environments" that influence society and culture rather than acting as direct shapers.

In this respect, Bruce Bimber (2003) offers an illuminating perspective on the relationship between democracies and technology. He relates each wave of information revolution with specific transformations in American democracy, particularly regarding shifts in party structures. For instance, during the first information revolution (1820–30), the postal system facilitated the mass (and national) dissemination of political information through newspapers, with parties holding sway as dominant entities characterized by hierarchical and centralized organization as well as national scope. In the second information revolution (1880–1910), industrialization led to the specialization and complexity of political information, giving rise to interest groups managing specific interests. The third information revolution (1950–70) saw the emergence of electronic media broadcasting information nationwide, primarily driven by commercial interests, thereby diminishing the monopoly of parties in campaign organization. In the ongoing fourth information revolution (1990–), marked by abundance and fragmentation, we witness the rise of postbureaucratic organizations shaping political landscapes.

The current context, characterized by the diffusion of artificial intelligence and big data,[1] can be considered a continuation, but at the same time it differs from the fourth industrial revolution analyzed by Bimber. AI focuses on advanced data processing and emulation of human intelligence, and it presents more complex ethical and governance challenges compared to the mere digitization of industrial processes. Therefore, we can ask ourselves: What is the impact of the arrival of artificial intelligence on democracy? Are we facing the rise of a sort of "artificial democracy"?

In his 1989 book *Democracy and Its Critics*, Robert A. Dahl outlines the defining characteristics of a polyarchy, the form that democracy has taken in the twentieth century (233). In a polyarchy, elected officials constitutionally hold control over governmental decisions regarding policy. In turn, elected officials are chosen and removed in relatively frequent, fair, and free elections. All adults have the right to participate in these elections, and most adults also have the right to stand as candidates for public offices. Citizens enjoy effectively enforced freedoms of expression, particularly in political contexts, enabling criticism of officials; governmental conduct; prevailing political, economic, and social systems; and dominant ideologies. Access to alternative sources of information, free from government or monopolistic control, is available. Furthermore, citizens possess effectively enforced rights to establish and join autonomous associations, including political parties and interest groups, which seek to influence the government through electoral competition and peaceful means.

At a time when the development of artificial intelligence is at breakneck pace and democracies are said to be in crisis, the expression "artificial democracy" – the title of this book – is intended to be a provocative one.[2] But what do we mean by artificial democracy?

The term "artificial" is rich with nuanced meanings that extend beyond mere fabrication. One interpretation pertains to human creation. This connotation suggests that something is crafted, manufactured, or devised by human hands rather than being naturally occurring. For instance, this is the case of the distinction between artificial and natural food production methods. While natural foods derive directly from the environment, artificial foods are synthesized or processed in laboratories, often involving chemical interventions.

Additionally, “artificial” can imply simulation or imitation of the natural world. In this sense, something may be artificially created to mimic a natural counterpart, despite lacking its inherent qualities. Take artificial flowers as an example. These replicas, crafted from synthetic materials, emulate the appearance of real flowers but lack the organic essence and vitality of their natural counterparts.

Moreover, “artificial” may connote falsity or lack of authenticity. This interpretation suggests that something is not genuine or sincere but rather contrived or insincere. Consider the notion of an artificial smile: a forced or feigned expression devoid of genuine emotion. While it may serve social purposes, such as conveying politeness or friendliness, it lacks the authenticity and warmth of a genuine smile.

Undoubtedly, democracy is a human-made system, thus it is “artificial” in the sense that it has been designed and implemented by humans rather than being a natural phenomenon. As for the other two meanings (artificial in the sense of “imitated” and in the sense of “false”), we can envision artificial democracy in two distinct ways.

First, democracy may be artificial because it is “conditioned” by AI and big data. AI and big data form the new environment in which democracies develop. For instance, they influence the functioning of democracy by affecting access to alternative sources of information and altering the forms of political participation (for instance, via organizations such as political parties), which are two among the main defining characteristics of polyarchies.

As regards the presence of alternative forms of information, content filtering algorithms on social media platforms may limit the diversity of opinions presented to users, thus shaping their political thinking and participation in the democratic process. On Facebook or X (formerly Twitter), we get more information in silos. Artificial intelligence reads our preferences and chooses the content that will be presented to us. We know that free elections also mean space for deliberation and debate to reach political compromises. However, in many liberal democracies, the opposite is happening online, with algorithms increasing polarization.

At the same time, political parties, on which representative democracy relies, have also changed their ways. Computers are replacing activists. Programmers replace traditional political organizers. With big data, they store personal information and target voters, making electoral behaviour even more malleable. With more refined technologies that promote microtargeting, it is becoming easier to develop fields of activation to build loyalty and modify citizen's values and attitudes. This transforms the essence of activism and political participation.

Second, democracy may be "manipulated" through AI and big data. The use of AI and big data impacts democracy by manipulating information, public opinions, and decision-making processes. This reflects the worries of one of the fathers of artificial intelligence, Yoshua Bengio, who has been taking this message to different forums, including the US Congress: "Democracy is not safe in an AI world." (Al Jazeera, 2023).

For example, AI and big data can be employed to manipulate elections through voter profiling, targeted dissemination of misinformation, and manipulation of survey results. With new technological tools combined with artificial intelligence, the level of control of the state over its citizens, rather than of citizens over the state, is likely to be greater. For instance, the digital tools developed in connection with COVID-19 have worried many.

With regard to the manipulation of elections, with Brexit, the referendum that allowed the United Kingdom to leave the European Union, the positive perception maintained by most media towards digital technologies has shifted from cool to suspicious. Following the March 2018 revelations of the Cambridge Analytica affair, political parties were now facing a large transparency deficit (Montigny, Dubois, and Giasson, 2019). A broader public then took the full measure of the power of big data and artificial intelligence to influence the electorate and the growing importance of these technologies in the conduct of election campaigns.

The influence of disinformation was starkly evident on January 6, 2021, when, spurred by online mobilizations within "filter bubbles" (Pariser, 2011), an unprecedented event unfolded in the United States. In what is regarded as the bastion of democracy, self-organized online supporters of Donald Trump, who refused to accept his electoral defeat, stormed the Capitol in Washington, DC. In another example, artificial intelligence now allows the manipulation of images, audio, and video clips, and we see and hear people saying things they never said. For example, during the 2024 New Hampshire primary, Joe Biden's voice was used to broadcast a fake message via robocall to tell people not to vote.

However, ultimately, we could also envision artificial democracy as a democracy "optimized" by AI and big data. Instead of having a negative impact, we could also consider how AI and big data can be used to enhance democracy. This might include secure electronic voting systems, predictive analytics to identify potential issues or inefficiencies in democratic processes, or the creation of online platforms that foster broader and more informed citizen engagement and empowerment and allow parliaments to evolve and open to the participation of constituents (Montigny and Bennan, 2020).

In any case, these matters represent a growing concern to members of the scientific and international communities. The United Nations Educational, Scientific and Cultural Organization (UNESCO) argues for a human rights approach to AI (UNESCO, 2024). A book published in 2023 by UNESCO and the Mila Research Centre of the University of Montreal concludes that AI offers great opportunities but also poses risks that were unforeseen just a few decades ago. The authors call for AI governance to be a global priority, mobilizing universities, governments, civil society, and international organizations (Prud'homme et al., 2023).

As regards the academic community, Andreas Jungherr (2023) argues for the development of a genuine research agenda on these topics. He sees four areas of concern for democracy in the age of AI. First, he notes its effects on self-rule. He suggests that "AI impacts both the ability of people to achieve self-rule and the perceived superiority of distributed decision-making over expert rule in complex social systems" (3). In short, algorithms make choices for individuals and filter information for them, with all the biases that technology can offer.

Second, Jungherr (2023) states that AI is a threat to equality. By predicting behaviour with observations from the past, AI "risks reinforcing existing biases in society and even porting socially, legally, and politically discontinued discriminatory patterns" (7). Thus, technology, based on the most abundant and available data, runs the risk of reinforcing certain prejudices to the detriment of minority groups.

Third, Jungherr (2023) also fears that "the public may come to believe that AI is actually able to offset the 'organized uncertainty' of democratic elections," thus weakening public trust in elections and acceptance of election results (8). By making elections more and more predictable, we can see them as a technical exercise or a show for which the script is written in advance, rather than a moment of deliberation and debate where the results can be surprising.

Finally, Jungherr (2023) concludes that "AI also affects the relationship between democracy and other systems of governance, such as autocracy, which some have argued has an advantage in the development and deployment of AI" (8). In authoritarian regimes, there are no limits to the use of technology to restrict civil rights. Surveillance innovations are developing rapidly. To what extent could democratic regimes be tempted to use these technologies in the interest of security and efficiency?

Evidence show us that established democracies are being transformed under the influence of big data and artificial intelligence. This encourages us to question ourselves more about the positive and negative effects of digital technology in our democratic life, as well as the risks it poses on things we take for granted.

An Integrated and Interdisciplinary Perspective

Over the past decades, several studies have examined the transformations of contemporary democracies. At the same time, we have witnessed a growing interest in technology-related issues within social sciences. However, research that considers the two phenomena simultaneously remains rare (Bigo, Isin, and Ruppert, 2019; Macnish and Galliott, 2020). What seems to be missing is an integrated and interdisciplinary perspective on the relationship between democracy and data. Similarly, since data is the main raw material for the development of artificial intelligence, the perspective we propose helps us to better understand the issues that are emerging in connection with a democracy that could become increasingly artificial.

With the help of several contributions and an original framework, this book analyzes the multiple relationships between democracy and digital data. To try to answer this question, we have mobilized works from different disciplines. This book is therefore a multidisciplinary collective work based on diversified approaches belonging to philosophy, law, sociology, communication, computer science, and political science. Our aim is therefore to contribute to this research field by adopting an approach based on three different but interconnected angles. They correspond to the three main dimensions of contemporary democratic politics: polity, politics, and policy.

Both electoral campaigns and public policies can be examined from the point of view of data's and AI's public regulation; that is to say, of the structure, legal and ethical, that forms and structures a political community (*polity*). Public data regulation poses dilemmas for which an interdisciplinary reflection on the role of the state is necessary (Haggart, Tusikov, and Scholte, 2021). It is the same for its consequences for the democratic process, as well as for ethical questions (Richterich, 2018).

We also know that in our highly connected societies, citizens' personal data are used by different types of political actors with different motivations. About *politics* – defined as the decision-making processes, the competition for power and participation – on the one hand, data and AI are used by political parties to profile citizens and send them personalized messages during election campaigns. On the other hand, citizens, activists, and party members could try to resist the so-called datification of politics. We will therefore analyze examples of these phenomena in electoral campaigns and implications for political actors, citizens, and the democratic system.

Finally, we know that data are used by governments, democratic or not, for surveillance (Lyon, 2014) and social control purposes (Završnik, 2017). The pandemic, with tracing applications and vaccination certifications or immunity passports, represents an important turning point in this respect, which deserves to be explored (French and Monahan, 2020). With regard thus to public policies (*policy*), the example of the pandemic will be used as a litmus test of the use of data in the concrete decisions of public actors and as an example of digital control by governments.

If the relationship between democratic regimes and data has been analyzed so far by focusing on a specific aspect (the dynamics of electoral competition, certain public policies, public regulation), this book integrates the three dimensions of contemporary politics and democratic regimes: polity, politics, and policy. The book is accordingly divided into three distinct sections. The first three chapters address the polity dimension in relation to the regulatory framework for a democratic use of data. The second section gathers four chapters related to the use of digital data by political actors and its multiple resistances. And the third section, with two chapters, specifically addresses the integration of digital data and surveillance in democratic states during the COVID crisis. In the next few pages, we will present each of these sections and the authors' contributions.

Polity: A Regulatory Framework for a Democratic Use of Data and AI

What are the effects of data campaigning on democracy? Who controls the data and what rights do we have as citizens over our personal data? What is the data's legal status? Public regulation of the use of personal data and the role of the state are essential issues in the current context (Haggart, Tusikov, and Scholte, 2021; Loveluck, 2015), both in relation to health issues (Déziel, 2018) and of the regulation of parties and electoral campaigns (Witzleb, Paterson, and Richardson, 2019) in Europe (Dobber, Ó Fathaigh, and Zuiderveen Borgesius, 2019) and in Canada (Déziel, 2019). In 2016, the European Union adopted the General Data Protection Regulation, which strengthens citizens' data protection within the European Union. Similar debates are underway in other jurisdictions, demonstrating that data protection is now seen as a matter of fundamental rights (McDermott, 2017). This first section is dedicated to public regulation of the use of personal data, to ethical debates, and to the comparison between the solutions adopted in different countries and regions. Which measures could be put in place to limit the manipulation of citizens? Which rules must be respected by policy makers if they are not to endanger the democratic character of our societies? To try to answer these questions, we mobilize three different perspectives: political philosophy, law, and computer science. The impression is that polity is struggling to keep pace with technological innovation.

The first chapter, written by a political philosopher, questions the effect of digital democracy on deliberation. François Blais insists on the risk for contemporary democracy of segmenting *demos* with social media and big data. He states that, ideally, democracy should allow citizens to make the "right choices." To achieve this, however, several conditions are necessary, such as the widest possible access to sufficient and reliable information, the relative independence of voters, and the articulation of fair contradictory debates on factual and normative issues. He argues that confrontation of different points of view is essential to this epistemic quality of democracy and that political parties have an important role to play in this regard. This is even more the case when the electoral system rewards those who seek alliances and who, to do so, force changes inside their own party and among the electorate. The excessive targeting of voters is part of a whole different dynamic and can instead cause political parties to lock activists and citizens into echo chambers where the search for compromise is neither necessary nor even desirable. He notes that we should not hesitate to put in place measures to limit the more damaging uses of some manipulations (control of electoral expenses, better protection of private data, the obligation for political parties to account for the data available to them, etc.). This would force political parties to make real contributions to the public debate.

By questioning the status of big data in a democratic regime, law professor Pierre Trudel also widens the discussion. He pleads for a legal recognition that big data is not only a private resource but also a collective resource. He argues that data, as traces of activity in a connected world, are now among the main inputs to value-generating activities. With the advent of massive data processing, he advocates that our understanding of data as a resource requires a paradigm review. When data are used to measure mass phenomena, they should be considered a resource that is collective in nature. He argues that applying a legal framework based only on relationships with individuals is absurd. For this reason, big data must be regulated differently. This means redesigning legal mechanisms to determine the rights and obligations of individuals, businesses, and governments in activities generating value from data. Doing so will mean acknowledging that, legally, big data is a resource used to create value. Such an approach is required to structure the legal regulations attributing the rights and obligations of big data as a collective resource.

In the next chapter, François Pellegrini, a professor of computer science, pushes the reflection even further. By imagining a democratic workbook, he anticipates the worst excesses and proposes a framework to prevent them. He argues that the irruption of information technology in administrative data processing in the 1970s, following the era of office machinery that dominated the first half of the twentieth century, gave rise to many fears about the infringement of freedoms that could result from the unequalled processing power of computers. In Europe, attempts to create state-owned centralized databases, interconnecting an individual's administrative data through a single identifier given at birth, met strong opposition from civil servants who witnessed the harm caused by the misuse of administrative files during World War II. These concerns led to the creation of the first "Informatics and Freedoms" laws in the world, originally targeted toward state-operated files but extended to privately owned files with the democratization of these technologies. However, more than half a century later, he notes that the lessons of history have not been learned. The "technological solutionism" that currently prevails as a doctrine for the management of human societies leads populations, often under the spur of fear but also seduced by smooth "user experience," to welcome the deployment of increasingly intrusive systems. These systems are shaping the contours of a surveillance society that, on the very pretext of preserving them, is corroding the principles that underpin democratic regimes. It is therefore necessary to propose to all policy makers and designers of computerized administrative systems a "democratic specification." This document expresses, in the form of functional requirements, the rules that must be respected by them if we are not to endanger the democratic character of our societies.

Politics: Use of Data, Profiling, and Personalizing in Electioneering

In most democratic regimes, electioneering has changed significantly over the last few election cycles, as is the case in Canada (Marland and Giasson, 2022; Marland, Giasson, and Lennox Esselment, 2017). The US presidential elections in 2016 and 2020, as well as the referendum on Brexit in the United Kingdom, have also shown that through the collection of data, political actors can adapt their campaigns to different types of voters by providing them with personalized messages and orienting the campaign on crucial themes (Bodó, Helberger, and de Vreese, 2017). We know that platforms like Facebook have the power to influence the way we inform ourselves, exposing us to opinions closest to our values (what Pariser [2011] calls the "filter bubble"). This can have consequences on citizens' trust in politics and on the legitimacy of the democratic process (Zuiderveen Borgesius et al., 2018), as well as for privacy (Bennett and Lyon, 2019). However, we still have a limited understanding of the impacts of microtargeting on the electorate (Lavigne, 2020). While some express serious concerns about the effects of microtargeting on democracy (Gorton, 2016), others say the effects are modest (Baldwin-Philippi, 2017; Kalla and Broockman, 2018). While it is true that microtargeting can bring benefits to citizens, it has been shown that the use of personalized information can disrupt electoral competition and compromise citizens' privacy, posing a challenge to democracies (Zuiderveen Borgesius et al., 2018). Furthermore, adaptations and resistances of parties, party members, and activists to these novelties, and to their public regulation, are still to be examined in depth: a thorough understanding of these dynamics, especially through qualitative data, is still missing.

In his chapter, Colin J. Bennett uses a well-known political science theory to explain the strategies adopted by political parties in Canada to develop their digital targeting tools while avoiding any limitations. He notes that Canadian political parties are not generally covered by Canadian privacy legislation – at either the federal or provincial levels. Recent efforts to bring them under the umbrella of privacy law, and the oversight of Canada's Information and Privacy Commissioners, have been met with stiff and unified resistance. Canadian parties obviously operate in a highly competitive political environment, but they are also entrenched and prone to collectively defend their interests against regulators. Indeed, there is some literature that suggests they operate as a form of "cartel" and have colluded in the past to exclude new parties from obtaining official party status, to shape the provision of state financial subsidies to benefit their own interests, and to regulate ballot access. Can the resistance to privacy regulation be explained by this same "political party cartel" theory? This chapter reviews the multipronged strategy to bring political parties within the ambit of Canadian federal privacy law over the last decade. That strategy has focused on parliamentary pressure, litigation, public and media advocacy, and targeted use of complaints to different regulators. Yet, after ten years of pressure, political parties are the one category of organization in Canada over which individuals have few, if any, privacy rights. A deeper understanding of this resistance is key to rendering data-driven elections more transparent and to extending privacy rights to individuals when their data is processed by Canadian political parties.

Some researchers claim to be able to guess a citizen's vote simply from the brand of their car: a Prius or a pickup? (Hetherington and Weiler, 2018). Other work suggests that how we consume our coffee is predictive of political allegiance (DellaPosta et al., 2015; Mutz and Rao, 2018). In their chapter, political scientists Catherine Ouellet and Yannick Dufresne present Datagotchi, an experiment conducted during the 2022 Quebec election. This recent research suggests that everyday choices (e.g., material and cultural consumption, leisure activities), although seemingly trivial, are often indicative of deeper social and political behaviours (Cavazza and Corbetta, 2016; de Moor and Verhaegen, 2020; Fischer and Mattson, 2009; Stolle et al., 2005). The predictive and explanatory power of lifestyle on voter behaviour is important because political parties increasingly use such data for segmentation and electoral targeting. In a context marked by the progressive decline in the predictive power of conventional sociodemographic variables on voting, this chapter suggests that lifestyle is a marker of socialization and therefore indicative of more complex human behaviour. Focusing on the Quebec and Canadian contexts in particular, the authors explain the context in which these transformations are occurring, discuss the practical issues associated with the richness of these new data, and address the ethical issues that arise because lifestyle data often takes the form of digital traces. This chapter also aims to raise awareness of the richness of personal data and the issues raised in the political domain by the digitization of privacy.

As we have seen, technology and the willingness of political party leaders to engage in data-driven campaigns is real. However, on the ground, these changes could face the shock of reality. Based on an ethnographic study carried out during an electoral campaign in Italy, political scientist Cecilia Biancalana exposes the limits and obstacles shared by party activists for artificial democracy. With electoral campaigns becoming more and more data-driven, the voluntary work of members and activists is said to have been replaced with that of professional campaigners. The relationship between party members and citizens with an unmediated leader-electorate connection is also said to be carried out through social network sites. On the other hand, on-the-ground and grassroots campaigns, in which the fundamental factor is the human one, regained their importance, but differently. In this new kind of campaign, human interaction is meant to lead to the construction of databases, to profile voters and contact them. The aim of this chapter is to analyze the experience of Noi Siamo Torino (NST), an electoral campaign inspired by the US model based on volunteers' field mobilization and microtargeting that took place in Turin in 2016. Through data collected during direct observation of the campaign, the NST experience and its meaning is investigated with the aim of detecting the changes in practices and repertoires of campaigning and party membership brought by a data-driven campaign. Campaigns based on canvassing and microtargeting appear both as the reinvention and modernization of ancient practices and as a substitute for the classical function of party membership.

Policy: Surveillance and Data Protection during the Coronavirus Pandemic

Why does the state care about citizens' data? Some will tell you that they are central to the notion of the state. Is the census not the basis of state authority? In fact, the action of the state has always been based on the collection of the population's data (Foucault, 2007). However, as Edward Snowden's revelations confirmed, in the digital age, the capacity of states (democratic or not) to monitor populations has increased (Lyon, 2014). Furthermore, the pandemic expanded what has been defined as "surveillance capitalism" (Zuboff, 2019). Indeed, according to sociologist Evgenij Morozov (2011), the internet is not an inevitably democratizing force but can also be used to limit the rights of citizens. In 2020, the outbreak of the coronavirus raised new questions for established democracies. In connection with the health emergency, for the first time, states have encouraged the development of mobile applications to trace contacts and thus limit contagion (Morley et al., 2020). Then, digital vaccination certifications were created (Phelan, 2020). If this is done to protect citizens' health, the question also opens new perspectives for surveillance, as well as ethical and legal dilemmas. How to balance the right to privacy and the protection of public health? What are the implications of collecting and using this kind of data for democracies? The last next two chapters explore, with health data in perspective, the state's reinforced capacity to monitor public policy in the digital age and the potential risks of this surveillance on the democratic rights of citizens.

In his chapter, sociologist David Lyon identifies the challenges posed to the social contract after surveillance capitalism met the pandemic. He reminds us that data-dependent "solutions" to problems posed by the COVID-19 pandemic proliferated, globally, in 2020 and 2021. These range from digital contact tracing to the establishment of large-scale data platforms for modelling and monitoring the progress of the virus, with many apps, devices, and systems. He argues that this represents the largest surveillance surge to date, easily outstripping post-9/11 surveillance expansion. Part of the reason for this is the rapid rise of platform companies that either work in partnership with governments (think of the Google-Apple Exposure Notification tools) or in parallel with lockdowns and other restrictions to provide online services of many kinds (monitoring work-from-home employees, supporting online learning or shopping, and the gig economy). While many benefited from online communication during the pandemic, a rapidly expanding surveillance capitalism has offered fresh opportunities to increase its reach. In so doing, arguably, government and companies together reduce freedom and fairness for ordinary citizens in the name of emergency measures. This in turn profoundly challenges the contested concept of a social contract, in which citizens agree to certain restrictions on liberty for the sake of social goods such as security, or, since, public health. Lyon notes a new democratic difficulty: public-private agreements query the idea that citizens and the state are the only actors. If the myth of social contract is worth anything, it will require a reset. Government-and-business partnerships, on the one hand, and digitally disempowered citizens and consumers, on the other, now struggle to recognize each other, let alone to develop a meaningful modus vivendi.

In the last chapter, Pierre-Luc Déziel, professor of information technology law, questions the management of the data accumulated by governments during the pandemic. He analyzes the Canadian case where an app using Bluetooth, COVID Alert, was developed so that a person could be notified of possible COVID-19 exposure. He notes that the use of this app has allowed federal and provincial authorities to collect a vast amount of personal health information (PHI) on Canadian citizens. Although considerable attention has been paid to the impact of such applications on the right to privacy of their users, most legal analyses primarily focused on the components of the application that are managed by the federal government. However, it can be argued that the data management operations that are the most delicate from a privacy standpoint were delegated to the provincial health authorities. The responsibility to pair the individual medical diagnosis with the pseudonymous key generated by the application rested indeed with the provinces. If the federal authorities have generally been transparent in the way they handled PHI collected and used via COVID Alert, the provinces have been generally more opaque. Little information has been given to the public regarding the type of data governance structure the provinces have put in place to handle PHI they gathered through COVID Alert and protect the privacy of its users. This chapter seeks to fill in this gap by presenting the results of a research project geared at amassing this information, notably through access to information requests, and providing an in-depth legal assessment of the privacy impact of these provincial data governance structures.

***

As you will see in this book, artificial democracy offers an opportunity for some and a concern for others. Many even see a major risk for deliberative democracy and for the protection of individual rights. With the speed of integration of digital and massive data by states and political parties, there is a need to fully understand this phenomenon. In many jurisdictions, artificial intelligence has already become a mediator between people and institutions. Today, the latter are controlled by actors with their own interests. But artificial intelligence may well become a political actor in its own right. And the implications for democracy can be manifold.

To integrate the complexity of what is happening now, we have gathered in this book experts from different fields, and we have adopted a framework with three different angles: polity, politics, and policy. This is an original contribution, but also a call to do more. In their conclusion, Julia Rone and Cecilia Biancalana bring some critical perspectives on the main findings of the book, placing them into the framework developed in this introduction. They also explore what could structure a new research agenda. Artificial democracy concerns both citizens and researchers. Both actors and regulators.

Initiated within a partnership between the University of Lausanne and Université Laval, this book is the fruit of a collaboration of several months between various researchers who collaborated in a workshop, then in a colloquium held in Quebec City in spring 2022. We wish to thank them. We hope that the following pages will help to trigger a larger discussion on the future of democracy.

Notes

[1] In the realm of emerging technologies, it is essential to distinguish between artificial intelligence and big data. AI refers to the capability of machines to emulate human intelligence and perform complex tasks such as pattern recognition, learning, and reasoning. Conversely the expression "big data" refers to large and complex datasets that are difficult to process using traditional data processing methods, often characterized by their volume, velocity, and variety. Although distinct concepts, AI often relies on big data to function effectively, leveraging large datasets to train models and enhance performance. This symbiotic relationship between AI and big data is crucial for understanding how they are reshaping democracies, as explored in this book.

[2] Credit for this title goes to Mickael Temporão, who proposed it during a brainstorming session at the project's beginning.

References

Al Jazeera (2023). Yoshua Bengio: Democracy is not safe in an AI world. August 12. https://www.aljazeera.com/program/talk-to-al-jazeera/2023/8/12/yoshua-bengio-democracy-is-not-safe-in-an-ai-world.

Baldwin-Philippi, J. (2017). The myths of data-driven campaigning. *Political Communication*, *34*(4), 627–33.

Bennett, C.J., and Lyon, D. (2019). Data-driven elections: Implications and challenges for democratic societies. *Internet Policy Review*, *8*(4).

Bigo, D., Isin, E., and Ruppert, E. (2019). *Data politics: Worlds, subjects, rights*. Taylor and Francis.

Bimber, B. (2003). *Information and American democracy: Technology in the evolution of political power*. Cambridge University Press.

Bodó, B., Helberger, N., and de Vreese, C.H. (2017). Political micro-targeting: A Manchurian candidate or just a dark horse? *Internet Policy Review*, *6*(4), 1–13.

Cavazza, N., and Corbetta, P. (2016). The political meaning of dining out: Testing the link between lifestyle and political choice in Italy. *Rivista Italiana di Scienza Politica,* 46, 23–45.

Dahl, R.A. (1989). *Democracy and its critics*. Yale University Press.

de Moor, J., and Verhaegen, S. (2020). Gateway or getaway? Testing the link between lifestyle politics and other modes of political participation. *European Political Science Review*, *12*(1), 91–111.

DellaPosta, D., Shi, Y., and Macy, M. (2015). Why do liberals drink lattes? *American Journal of Sociology*, *120*(5), 1473–511.

DemDigest. (2022). World Values Survey confirms democracy's long-term but fragile appeal. *Democracy Digest,* December 16. https://www.demdigest.org/world-values-survey-confirms-democracys-fading-fragile-appeal.

Déziel, P. (2018). *La protection des renseignements personnels sur la santé au temps de la biosécurité*. LexisNexis.

Déziel, P. (2019). Microciblage politique et protection des renseignements personnels au Canada: mieux protéger la vie privée pour défendre l'espace public. In A. Bensamoun and F. Labarthe (eds.) *Culture et numérique*, 185–212. Éditions Mare et Martin.

Dobber, T., Ó Fathaigh, R., and Zuiderveen Borgesius, F. (2019). The regulation of online political micro-targeting in Europe. *Internet Policy Review*, *8*(4).

Fischer, C.S., and Mattson, G. (2009). Is America fragmenting? *Annual Review of Sociology*, *35*, 435–55.

Foucault, M. (2007). *Security, territory, population: Lectures at the Collège de France, 1977–78*. Springer.

French, M., and Monahan, T. (2020). Disease surveillance: How might surveillance studies address COVID-19? *Surveillance and Society*, *18*(1), 1–11.

Gorton, W.A. (2016). Manipulating citizens: How political campaigns' use of behavioral social science harms democracy. *New Political Science*, *38*(1), 61–80.

Haggart, B., Tusikov, N., and Scholte, J.A. (Eds.). (2021). *Power and authority in internet governance: Return of the state?* Routledge.

Hetherington, M., and Weiler, J. (2018). *Prius or pickup? How the answers to four simple questions explain America's great divide.* Mariner Books.

International Institute for Democracy and Electoral Assistance. (2022). *Global state of democracy report 2022: Forging social contracts in a time of discontent*.

Jungherr, A. (2023). Artificial intelligence and democracy: A conceptual framework. *Social media+Society*, July–September, 1–14.

Kalla, J.L., and Broockman, D.E. (2018). The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. *American Political Science Review*, *112*(1), 148–66.

Lavigne, M. (2020). Strengthening ties: The influence of microtargeting on partisan attitudes and the vote. *Party Politics*, online first.

Loveluck, B. (2015). Internet, une société contre l'État : Libéralisme informationnel et économies politiques de l'auto-organisation en régime numérique. *Réseaux*, *192*, 235–70.

Lyon, D. (2014). Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data and Society*, *1*(2), 1–13.

Macnish, K., and Galliott, J. (2020). *Big data and democracy*. Edinburgh University Press.

Marland, A., and Giasson, T. (2022). *Inside the local campaign constituency: Elections in Canada.* University of British Columbia Press.

Marland, A., Giasson, T., and Lennox Esselment, A. (2017). *Permanent campaigning in Canada.* University of British Columbia Press.

McDermott, Y. (2017). Conceptualising the right to data protection in an era of big data. *Big Data and Society*, *4*(1), 1–7.

Montigny, E., and Brennan, A. (2020). Pétitions électroniques au Québec : entre transfert et résistance, *Participations*, *28*(3), 151–76.

Montigny, E., Dubois, P., and Giasson, T. (2019). On the edge of glory (... or catastrophe): Regulation, transparency and party democracy in data-driven campaigning in Québec. *Internet Policy Review*, *8*(4).

Morley, J., Cowls, J., Taddeo, M., and Floridi, L. (2020). Ethical guidelines for COVID-19 tracing apps. *Nature*, *582*, 29–31.

Morozov, E. (2011). *The net delusion: How not to liberate the world*. Penguin.

Mutz, D.C., and Rao, J.S. (2018). The real reason liberals drink lattes. *PS: Political Science and Politics*, *51*(4), 762–67.

Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin.

Phelan, A.L. (2020). COVID-19 immunity passports and vaccination certificates: Scientific, equitable, and legal challenges. *The Lancet*, *395*(10237), 1595–598.

Prud'homme, B., Régisand, C., and Golnoosh, F. (Eds.). (2023). *Angles morts de la gouvernance de l'IA*. UNESCO / Mila – Institut québécois d'intelligence artificielle.

Richterich, A. (2018). *The big data agenda: Data ethics and critical data studies*. University of Westminster Press.

Stolle, D., Hooghe, M., and Micheletti, M. (2005). Politics in the supermarket: Political consumerism as a form of political participation. *International political science review*, *26*(3), 245–69.

UNESCO (2024). Artificial Intelligence: Rapid technological advancements in AI are transforming disciplines, economies, and industries, and challenging ideas about what it means to be human. See: https://www.unesco.org/en/artificial-intelligence.

Witzleb, N., Paterson, M., and Richardson, J. (Eds.). (2019). *Big data, political campaigning and the law: Democracy and privacy in the age of micro-targeting*. Routledge.

Završnik, A. (Ed.). (2017). *Big data, crime and social control*. Routledge.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.

Zuiderveen Borgesius, F., Möller, J., Kruikemeier, S., Ó Fathaigh, R., Irion, K., Dobber, T., Balazs, B., and de Vreese, C.H. (2018). Online political microtargeting: Promises and threats for democracy. *Utrecht Law Review*, *14*(1), 82–96.

PART 1

Polity: A Regulatory Framework for a Democratic Use of Data and AI

This page intentionally left blank

1 Big Data and Electoral Democracy
The Epistemic Risk

François Blais

Social Media, Democracy, and Politics

Political parties use social media for two different types of communications: open communications like traditional forms in contemporary mass media (newspapers, television, or radio) and closed communications and personalized messages for specific groups through the internet.

These two types of communications measure different challenges to the quality of democratic life. However, this chapter focuses on closed communications, a phenomenon that is less popular among political scientists than open communications, which remain at the heart of partisan political activity, especially during election periods. We all know about some communication strategies that use negative advertisements in an attempt to seduce the electorate by simply demonizing the adversary (Maarek, 2011). The risks of these very partisan communications that appeal to negative emotions have been well known and called out for a long time even as they remain a nuisance for the quality of democratic debates. Closed communications, which have received less attention and thought from political theorists (Lees-Marshment, 2009), also represent risks for democracy, but they are risks of a completely different nature, and they are worth understanding if we want to control or reduce them.

In contemporary democracies, closed communications are made possible by the harvesting of millions of personal data points collected on social networks by private agencies and sold to political parties for their partisan goals (to join, convince, and mobilize). Data are processed using powerful algorithms that make it possible to define profiles of interests or "vulnerabilities." Online political microtargeting adapts its content to catch the eye of potential voters and eventually get their votes. In a hardly exaggerated way, one could say that political parties address each voter as a close friend who sometimes knows you better than you know yourself. Thanks to big data, political parties are often aware of one's intimate interests and undisclosed concerns. This allows them to address a potential voter by establishing a close relationship from the start, which is a great advantage when it comes to getting one's attention. They can choose the political topics that may interest you personally and find the right words to present them to you. It also allows them – and this is from a strategic point of view, which is just as important – to avoid presenting elements of their platform to which the voter does not grant the same importance or which they may even be against. The goal of political parties is therefore not so much to inform as to convince by providing arguments that will please the potential voter.

These techniques of profiling and targeting messages are well known in the field of commercial marketing (Pariser, 2011), but in politics, their use is recent and still uneven across jurisdictions (Macnish and Galliott, 2020). They were the subject of more media attention following Barack Obama's 2008 campaign ("the first social media president") and after the British referendum campaign that led to Brexit in 2020. Grassegger and Krogerus (2017), for example, claimed that microtargeting helped Trump's 2016 victory in introducing a new type of political "neuromarketing" (Hegazy, 2019). Actually, one of the most controversial players in this campaign was Cambridge Analytica, an important data analytics company that used data from 50 million profiles to design political messages for different customers. According to Frier (2018), Donald Trump's team invested 44 million dollars and Hilary Clinton's team 28 million dollars in digital advertising during the 2016 presidential campaign. Since that time, the use of digital marketing for political purposes has expanded, but it remains difficult to accurately establish its use since political parties are reluctant to provide more information (see Chester and Montgomery, 2017). It is clear, however, that this practice has become more important in the United States given the sums of money invested in campaigns, exceeding a billion dollars for a presidential election, and the increasingly important relative share for digital spending (Baum, 2020). This practice also tends to settle on the European continent but in obviously less significant proportions (Dommett, 2019). For example, during the German federal election campaign in 2017 and the European election campaign in 2019, German political parties invested in digital advertising on Facebook and Google to the sum of more than 4 million euros according to evaluations carried out by Hegelich and Serrano (2019).

Of course, targeting and clientelism are not new in politics. Politicians have always tried to influence voters by adjusting their message and emphasizing what a voter wants to hear. Politicians know that it is necessary to understand what may interest voters if they are to have the right tone and formulate the political proposal to seduce them. This political reality is as old as democracy itself if we stick to the writings of Plato on the subject! The novelty here lies, first, in the precision of the profiles that algorithms and AI allow (tastes, concerns, interests) and, second, in the number of potential voters potentially reached by these new technologies. Never in the past has it been possible to achieve what can now be done with social networks and online political microtargeting (Gorton, 2016). We can even go so far as to say that, with algorithms, it is now possible to determine whom you will probably vote for in the next election without you even knowing it yet (see Ouellet and Dufresne, this volume; Theviot, 2019). As these techniques of online microtargeting will most likely see a decrease in costs over time and, therefore, a certain ease in their expansion within democracies, taking more interest from now on in trying to assess their consequences for democratic life is justified.

Political parties and providers of personal data defend personalizing by claiming that it makes it possible to reach more voters at a lower cost, to inform them about more singular elements of their platforms, which should interest them, and to increase political participation. Its strategic value is very important for them. In a context where militancy within political parties is on the decline in most well-established democracies and a growing number of voters no longer use traditional media, political parties will increasingly use social media to reach them and to recruit volunteers and donors (see Biancalana, this volume). That is why it is important to dwell on this subject and examine some potentially harmful consequences even though it remains yet unknown if these new methods can completely alter the outcome of an election.[1]

In fact, my interest here is not in the political efficacy of online microtargeting to purportedly win or "steal" an election. This issue exists, experts disagree on it, and recent results from empirical investigations are not decisive.[2] However, I do not believe that this represents the most critical danger for democracy, especially if we protect a competitive electoral context between political parties. In fact, if all major political parties have access to similar resources to carry out their electoral campaigns, I personally think that their relative advantage in using these new techniques, if they are even effective and can really make the difference between winning or losing an election, will also be similar. Therefore, this is not the question I want to discuss in this chapter.

I am more concerned about the quality of democratic life if, one day, microtargeting replaces open communication in democracy to the point that we no longer foster genuine debates during election campaigns. As for the debates themselves, recent trends have not necessarily displayed any improvement in their quality, quite the contrary, and the use of closed communications should ideally not accelerate a decline in these debates. But to demonstrate this risk, I need to define beforehand what a democracy is and, more precisely, what we can (and ought to) expect from it.[3] My subject is therefore essentially on the normative side of political theory. I do not seek to explain a new political phenomenon but rather to elaborate arguments on its desirable or undesirable character (for an introduction to normative issues and democracy, see Cunningham [2002]).

Procedural Democracy versus Epistemic Democracy

To know if a democracy is adequately performing, it is first necessary to define a standard: what one can reasonably wish and try to improve in it. Proponents of a more "procedural" conception of democracy will generally need the following three requirements: regular and competitive elections, universal suffrage, and respect for some fundamental freedoms.

These elements are of course very important. The first advantage is that they are quite operational without presupposing any expected political results (Przeworski, 2010). Historically, these arrangements have facilitated the peaceful resolution of many potential conflicts and the establishment of responsive governments (Manin, 1997). It took centuries of fighting to get these institutions in place, as well as the many other secondary arrangements that make it possible to be effective (separation of powers, parliamentary rules, support for political parties, etc.). Even today, billions of people are not fortunate enough to live under this form of responsive government, and several democracies that were believed to be well established could be in danger (Mounk, 2019). Therefore, the conditions mentioned are necessary for the definition of modern democracy, and a consensus among all the major international organizations responsible for assessing the state of democracies in the world, such as the Economist Index of Democracy, must be achieved. This does not mean that these conditions are sufficient. It is indeed possible to have somewhat more demanding and legitimate expectations. Let us see why, taking as an example the issue of closed political communications.

If we limit ourselves to these requirements to define democracy, we can conclude that the personalization of partisan information by political parties, for example, does not appear to constitute a great risk: the personalization of messages does not deprive anyone of the right to vote and does not prevent us, a priori, from having competitive elections. This is even more true if political parties have access to financial resources that allow all of them to use the same technology. In no way, one might add, could this strategy of partisan communication be considered detrimental to the exercise of fundamental freedoms within any modern democracy. Thus, according to the three elements selected above, we would not have to worry about political parties' increased use of these online microtargeted communication because the competitive character of the elections is protected. This poses a serious problem: democracy is not only about fair competition for power, as some minimalist theorists like to affirm.

Thus, those three requirements, despite their historical and political importance, are not yet sufficient to justify democracy. An important element is still missing. Ideally, the democratic procedure presented above should help a community make “the right choices” or well-informed decisions. This idea is very old and goes back in fact to Aristotle, in *The Politics*, who pointed out that involving many people in a decision was generally superior to involving only the most brilliant of them. This dimension of democracy is called today the epistemic (or cognitive) value of democracy. It is important but also controversial. Many theorists even believe that it is nonexistent and that nothing should be expected from it. Others, like me, argue that it remains necessary to assess the quality of the democratic experience to subject democracy to a more robust finality (see Estlund, 2007). But what can it mean to make the “right choices” in a democracy? This is based on two complementary procedural requirements: choices supported by rational and informed arguments or choices supported by impartial rather than self-centred arguments.

The first requirement is easy to understand: the choices we make should be as informed as possible so their realization corresponds to what we are really looking for and to avoid any form of deception or misunderstandings. This means that we should be able to share our knowledge of the facts surrounding a decision. This seems obvious even if it is not always easy, and it is often far from it. It often happens that there remain disagreements on the facts themselves or on the consequences that can be expected from one political choice rather than another. This limit is specific to the human condition and to our uncertainty about the future and knowledge. Nothing can be changed about it. This difficulty being well known, it remains necessary to try to reduce its risks by favouring as much as possible the use of rigorous arguments that give way to facts – to reason more than to emotion or passion. That is what I mean by choices supported by rational and informed arguments. In politics more than anywhere else, we should always favour rational debates even if, obviously, the opposite too often occurs in a partisan context. Our decisions will be more thoughtful and their consequences more predictable.

The second requirement stems from the fact that our political choices necessarily have consequences for a multitude of people. The morality in these circumstances is to favour policies that do not unduly harm others. This requires some form of decentring or impartiality in the way we justify our political choices to others. We cannot, for example, defend a political choice on the sole basis that it is in our sole personal interest. On the contrary, decentring requires consideration of the desires and interests of others. Not to the point of forgetting oneself but placing as much value on what others may want or think (Beerbohm, 2012). It is important to emphasize that this perspective does not necessarily lead to immobility but to the search for consensus when they exist. To achieve this, the others must also have the same concern for impartiality and decentring. Otherwise, they will at least have to explain the reasons for not recognizing in others the right they grant to themselves. But is this requirement of impartiality realistic? Neither more nor less than that of rationality above. In both cases, these are criteria for evaluating the quality of our public debates and the epistemic consequences they may entail. Perfection is therefore not required. What counts are the potential benefits that can be drawn from the application, even marginally, of these constraints.

These two requirements, rationality and impartiality, are logically distinct but linked in practice: if "the right choices" were well-informed choices, it would mean that a group could knowingly favour actions that are very harmful to another group. Moreover, if "the right choices" correspond only to moral arguments without concerns about feasibility, this could allow the implementation of policies that do not have any serious basis and are therefore doomed to failure. The democratic ideal requires well-informed citizens seeking the common good. Of course, I am speaking here of an ideal. These requirements are not necessary to understand how democracies work, but they are useful to determine how to improve them. But to do this, one must first answer the following question: Are these new requirements realistic or even desirable in making "the right choices"?

The Limits of Electoral Democracy for an Epistemic Approach

The idea that democracy, in particular electoral democracy, can help us in making "the right choices" is enough to put off many observers of political life. In fact, there is a vast literature in political science, on opinion polls carried out for decades, that invariably concludes by underlining the important limits of contemporary average voters. These limits have direct consequences for the quality of our democratic choices.

Some will even go so far as to conclude that they call into question the legitimacy of electoral democracies and ask for radical reforms (Brennan, 2016; Caplan, 2007). They are based on the following empirical findings: voters are poorly informed if not incompetent in political matters, voters are unable to name political leaders or speak properly about major public policies, and most voters are also unable to situate a political party on a left-right axis or to identify the main orientations of the most important political parties.

These findings are not the result of cynical prejudices; they have been observed and measured for decades, particularly in the United States but also in other countries. Yet voters are no dumber than in the past. The vast majority are even much more educated. This situation is mainly explained by the lack of interest in the population for politics in general and especially by the time-costs necessary for one to correctly understand them. There are of course exceptions, but that is not enough to counterbalance the relative ignorance of most voters towards politics according to surveys (Achen and Bartels, 2016). Moreover, being a little more engaged does not necessarily protect one from some forms of irrationality. Surveys show that "aware voters" are too often biased: they overestimate the advantages of their political orientations and underestimate the disadvantages, they systematically favour evidence that reinforces their opinions, and they tend to associate with one group and demonize other groups (Brennan and Landemore, 2022).

Ignorant in political matters or blinded by its biases, this is the sad image that emerges of the average voter in contemporary democracies. So, if democracy is valuable, conclude the American political scientists Achen and Bartels (2016) in their book *Democracy for Realists: Why Elections Do Not Produce Responsive Government*, a book that lists all the research carried out on the "skills" of voters for fifty years, it is not because of the individual skills of voters in the matter. It is rather because it promotes political alternation through competitive elections and protects a certain number of fundamental freedoms, including freedom of expression. To use the famous formula of Joseph Schumpeter, founder of the realist school of democratic theory, the only advantage of democratic regimes is their ability to "throw the rascals out." Why expect any more at the risk of being disappointed?

Should we be surprised by this sad observation? Not really. Is it true, consequently, that empirical observation supports, up to a certain point, the more "realistic" or "minimalist" conceptions of democracy over the defenders of the more epistemic conception? Not necessarily. Democracies are human institutions, not ideal institutions. For those reasons, they will never be immune to mistakes or irrational and selfish decisions. But if all this is true, how is it possible to continue defending the epistemic potential of democracy? The correct answer is that it offers a perspective to improve the democratic experience while remaining very lucid about our expectations.

It must be remembered that the theoretical discussions on the subject take place at two very different levels: empirical theory versus normative theory. We must try not only to understand how electoral democracy works but also to establish on what basis we can not only justify but above all improve it. It is on this last point that a minimalist and purely procedural conception of democracy appears insufficient in my point of view. It is quite possible to admit the many limitations and flaws of electoral democracies while indicating the horizon by which they can realistically be improved. The error of some enthusiastic supporters of the so-called realist school of democracy is to have too often confused what "is" and what "should be."

We must never deny the realities of electoral misinformation or bias; on the contrary, we must be sufficiently aware to limit their most damaging effects. But because their epistemic potential is an attribute that must be valued if we do not want to settle for a purely procedural design, we must make sure to find more optimal institutional arrangements when possible. At the heart of this research, we must find the requirements of rationality and impartiality, as various defenders of the epistemic approach have sought to do in recent years.[4]

Improve Democracy by Deliberation

There are two complementary ways to improve the epistemic quality of democracy: improving the average level of competence of active voters or improving the deliberative capacity of political actors and those who follow them. These approaches are, under the right conditions, mutually reinforcing. Of course, the first approach is the most attractive, but it probably remains the most difficult to achieve in mass democracy since it raises practical and moral issues.

According to Achen and Bartels (2016) and economic theories of democracy, the main obstacles to active voters' political competence is not education but rather the time necessary to devote to these questions and the control of personal biases that can hinder the consideration of the complexity of an issue presented to the public. How does one ensure, in a free and modern but also very individualistic society, that citizens have more interest in public affairs? That they learn more about major political issues? That they invest the necessary time to be properly informed and openly discuss major political issues with those close to them? One can wish it, of course, but is it realistic to require it? I do not think so.

Given this situation, a radical solution would be to restrict the right to vote only to people who have demonstrated a minimum mastery of the issues before voting, as some contemporary authors suggest (see Brennan [2016] or Caplan [2007] for some proposals to limit or regulate the right to vote). But this solution poses insurmountable practical difficulties, first, and it is unacceptable from the moral point of view. It would lead, among other things, to taking the right to vote away from many voters, many of whom are found in the most disadvantaged strata of society where there is already a high level of abstention (Estlund, 2007). This proposal also denies the principle of equality between citizens and the universality of the right to vote, one of the great achievements of contemporary representative democracies. That said, developing a new ethics of voting in which this practice is no longer considered an individual right but above all a collective responsibility remains, in our time, a priority that is perfectly in line with the ideals of epistemic democracy. Much remains to be done on this subject to change perceptions.

Strengthening the deliberative capacity of political actors, the second strategy presented above, is probably a more realistic one, but it also encounters many obstacles. The search for the truth or compromise is not, as we know, central to political activities, especially if it involves partisan activities. For politicians, politics is more about strategy and persuasion. Authentic deliberation is not an essential key to winning an election. Sometimes, the opposite is true.

For these reasons, some would argue that it could be necessary to take certain political issues out of the hands of politicians and hand them over to a group of usually randomly selected "ordinary citizens" (see Landemore, 2020). They would have to debate these issues in a forum that is sometimes called a "citizen assembly," where the level of information will be optimal and the biases minimized since they are not fuelled by partisanship. But this strategy, which essentially emphasizes the quality of the deliberation of a very small group of people, can never fully replace modern democracy, which alone has the necessary legitimacy to elect representatives and let them govern effectively (see Goodin [2008] for a survey). In fact, these "citizen assemblies" can at best support elected officials in the examination of complex questions whose treatment requires less partisanship (see Landemore [2020] for more examples to improve the deliberative aspect of democracy). This cannot replace a representative democracy made up of people chosen by an entire population following an election by universal suffrage.

Another way to improve the deliberative capacity within our democracies is "simply" to be more demanding of its main political actors, the political parties. This deliberative capacity must, however, be understood in a broad sense. It is not only a question of ensuring that political debates are more rational and better argued, which is particularly difficult in a partisan context. It is also necessary that the information used for decision making be as complete as possible and that politicians explain their orientations more publicly. Indeed, the quality of deliberation depends as much on the effort to share arguments as on the quality of the information available to voters. On this last point, political parties and politicians can do much more if they are better equipped to do so. I will return to this point in the conclusion.

From Echo Chamber to Public Reason

Targeted communications, in general, are not part of good practices meant to enlighten public reason. There are several reasons for this. Research indicates, first, that election candidates who make online targeted ads are likely to be much more divisive in their speech than in their overt ads for strategic motives (Goodman et al., 2017). This should not surprise us. The purpose of this type of communication is to reach people whose interests and, sometimes, political convictions we already know. We do not seek to convince the greatest number with all the compromises that are sometimes necessary to get there. On the contrary, we insist on very specific elements that are distinct from that which we sometimes give exaggerated importance. Social media networks have accustomed us to this sad reality for a long time. In this context, the natural tendency is to radicalize the political discourse since the opponents of this discourse cannot respond or simply present a political issue from another perspective.

Second, research also shows that restraining one's audience to a specific public facilitates exaggeration and the use of arguments that flatter the ego or that support the interests of voters rather than the common interest (Gorton, 2016). This, of course, is due to the highly personalized form of this communication. In the virtual absence of opponents, it is easier to exaggerate one's words, to caricature one's adversaries, to be kind with like-minded people, and to hide the truth. The competitive nature of political life and the functioning of social networks obviously accentuates this tendency, which is not in itself new to politics but which is magnified by the personal information harvested on social networks and by the use of AI and powerful algorithms.

Third, separate audiences also make it possible to communicate different and sometimes even perfectly contradictory information. As the public debate becomes fragmented between different groups and different subjects, a political campaign might even decide to completely ignore some voters it does not need to win. Again, this is nothing new in politics but, with the new tools provided by the digitization of personal data, it can become dangerously efficient (Wooley and Howard, 2017). Under these conditions, online political communications can become powerful tools for propaganda. The nonpublic nature of targeted communication closes the door to a minimal deliberative space where arguments can be evaluated on their merits. Instead of strengthening the critical thinking capabilities of citizens, it necessarily diminishes them since they are maintained in their beliefs. This profiling, let us not forget, is based on a level of highly personalized information at a level that has never been reached before in the history of democracies. They can even get you to agree to proposals that you had never thought of but that are related to your profile and your fields of interest rather than to your responsibility as a citizen.

To summarize, online microtargeting political communications tend to associate people who share the same interests. This is the main purpose of these closed communications. The risk is to confine users to long-term echo chambers, a situation in which various groups of citizens share and reinforce their own ideas and where contrary opinions have no chance to confront them (Kinkead and Douglas, 2020). If information is limited to what each group wants to hear, confirmation biases will be automatically reinforced. An echo chamber can be seen as a form of closed "epistemic community," centred around and defined by specific beliefs, which excludes and discredits different points of view. In such unilateral communicative situations, opposing positions risk being caricatured, demonized, or simply ignored. It is also much easier to play on emotions in this context, to the detriment of rationality or impartiality since no justification or explanation is required in this context.

This is why, to avoid this situation, public rather than private communications in political matters should be encouraged. These favour, without ever guaranteeing, of course, confrontation of arguments and establish conditions for a "free market of ideas" to quote John Stuart Mill (1859). Public debate makes it possible to detect falsehoods, correct errors, and continuously improve theories and practices. Advertising our communications also forces us to argue in a more reasonable way with a greater respect for facts since our arguments will inevitably be submitted to our adversaries but also, in contemporary democracies, to media scrutiny.

From a more philosophical perspective, the fundamental differences between public communications and political communications reserved for a specific group is the absence of "publicity." By "publicity," I mean a particular form of communication and argument between humans talking about politics and the common good. Publicity is not only a mode of open communication conducive to establishing the truth but also a form of argumentation necessary in a pluralist society (Rawls, 1993). Public reason, however, requires that we develop an argument that is addressed to as many people as possible beyond the circumstances of each voter. Many risks are involved if we multiply closed communications rather than public communications.

We owe the notion of "public reason" to the philosopher Habermas,[5] who defined the rules of successful public communication. These rules, such as freedom, equality, rationality, and publicity, were subsequently taken up, in one form or another, by proponents of the deliberative conception of democracy such as James Fishkin (2018). These rules require inclusivity and the presentation of any political argument in such a way that it is potentially possible for anyone wishing to reach an agreement to do so. A democratic society that respects this requirement has the great merit of making the application of majority rule more legitimate since this will not be based only on the number of supporters but above all on the value of the public debate that precedes political choices. The heart of moral legitimacy of democracy, from this perspective, is therefore less majority rule than publicity rule.

Conversely, online political microtargeting targets the personal interests of voters and is meant to get their attention and their votes but not to broaden the public debate. This implicitly sustains the idea that democracy boils down to the aggregation of individual preferences and the cool application of majority rule, which would have every right to impose its will on those who disagree with it. According to this unfortunately too widespread vision, the right to vote becomes an absolute right, which can be exercised without preconditions and without justification. Democracy, for its part, then becomes a game of power, of a balance of power, and loses all its epistemic potential.

Conclusion: The Epistemic Role of Political Parties

The debates surrounding the issue of the epistemic quality of democracy have occupied a lot of space in recent decades. While for centuries the main issue was that of equal voting rights, that of the epistemic quality of democracy has taken on a certain importance and this should be considered moral progress. The epistemic value of democracy is not the only element to consider when we examine the risks confronting contemporary democracies. Other values such as efficiency, legitimacy, impartiality, and accountability can also be invoked in the name of democracy. Therefore, I insist on the usefulness and importance of political parties for a healthy democracy – most of which, let us remember, within contemporary democracies benefit from public funds and have responsibilities to voters and people affected by their decisions.

Political parties are necessary to animate political life, to defend points of view on living together, and to propose a direction for a community. Most citizens, for lack of time or interest, are unable to develop such elaborate political thought. Political parties also have time, interest, and resources to devote to these issues and propose collective initiatives. They are made up of "professionals" of politics, practitioners capable of managing a modern public administration. Moreover, when party discipline is strong, their political action is more predictable, which increases their accountability and that of voters by incidence. Through them, a population can influence the course of events and the way they will be governed by the choices it makes. This is a major asset of democratic functioning. How it is possible to increase their epistemic potential any further is a question that remains to be determined.

It could be to implement institutions and practices that value epistemic potential and that force politicians to present their political decisions and defend them in the most rigorous way possible. This may, for example, involve having to respond frequently to a panel of citizens, experts, or journalists. This may also take the form of presenting their policy platforms in such a way as to facilitate critical examination by specialists or other political parties. Political parties should ideally make accessible, both in form and content, the analyses and justifications of their policies. The format chosen is important. If their decisions are based on reliable and intelligible data, they can more easily be verified and possibly challenged by opponents, experts, or activists. The epistemic potential of democracy will be enhanced simply by positive arrangements made by political parties. In these matters, there is really no limit to transparency of ideas and facts. And closed communications do not really play a useful role here.

My intention with this chapter is not to be alarmist by claiming, for example, that closed communications represent the greatest risk for contemporary democracies (see Pellegrini, this volume, for an original point of view about this issue). I have no empirical evidence that this is necessarily the case. But if ever we conclude that the damage of online microtargeting and increasingly segmented systems of electronic channels are significant, which is quite possible, the use of personal data by political parties should be restricted. It could be by stricter legislation on the use of personal data provided by social media companies, by the obligation for political parties to report on the use of data at their disposal, or through a corollary monitoring of allowable election expenses (Zuiderveen Borgesius et al., 2018). Some countries are stricter than others on this subject. We should do this in the name of improving public debate but also to favour the accountability of political parties, which have a responsibility for promoting public debate to help voters make "the right choices." For the many reasons that I have tried to explain in this chapter, there is little chance that online microtargeting can promote their epistemic roles in the future.

Notes

[1] See Kalla and Broockman (2018) for an extensive empirical study about the "persuasive effect" of campaigns in democracy. See also Baldwin-Philippi (2017) on some "myths" about data-driven campaigning. For recent surveys about this "efficiency" issue, see Kefford et al. (2022) and Tappin et al. (2023).

[2] "The extent to which this approach [i.e., microtargeting] confers a persuasive advantage over alternative strategies likely depends heavily on context" (Tappin et al., 2023, p. 1).

[3] In this text, I will address the issue of the consequences of microtargeting for democracy. Other related ethical issues like "data breach" or "privacy" are also important. See Bennett and Lyon (2019) and Stelzer and Veljanova (2020) on these important issues. See also Trudel, this volume, and Bennett, this volume.

[4] See Brennan and Landemore (2022) for two opposite solutions.

[5] See Habermas (2023) for a recent and clear discussion.

References

Achen, C.H., and Bartels, L.M. (2016). *Democracy for realists: Why elections do not produce responsive government.* Princeton University Press.

Baldwin-Philippi, J. (2017). The myths of data-driven campaigning. *Political Communication*, *34*(4), 627–33.

Baum, L. (2020, October 29). Presidential general election ad spending tops $1.5 billion. *Wesleyan Media Project.* https://mediaproject.wesleyan.edu/releases-102920/.

Beerbohm, E. (2012). *In our name: The ethics of democracy*. Princeton University Press.

Bennett, C.J., and Lyon, D. (2019). Data-driven elections: Implications and challenges for democratic societies. *Internet Policy Review*, *8*(4).

Brennan, J. (2016). *Against democracy.* Princeton University Press.

Brennan, J., and Landemore, H. (2022). *Debating democracy: Do we need more or less.* Oxford University Press.

Caplan, B. (2007). *The myth of rational voter: Why democracies choose bad policies*. Princeton University Press.

Chester, J., and Montgomery, K.C. (2017). The role of digital marketing in political campaigns. *Internet Policy Review*, *6*(4), 1–20.

Cunningham, F. (2002). *Theories of democracy.* Routledge.

Dommett, K. (2019). Data-driven political campaigns in practice: Understanding and regulating diverse data-driven campaigns. *Internet Policy Review*, *8*(4), 1–18.

Estlund, D. (2007). *Democratic authority: A philosophical structure.* Princeton University Press.

Fishkin, J. (2018). *Democracy when the people are thinking: Revitalizing our politics through public deliberation*. Oxford University Press.

Frier, S. (2018, April 3). Trump's campaign said it was better at Facebook. Facebook agrees. *Bloomberg*. https://www.bloomberg.com/news/articles/2018-04-03/trump-s-campaign-said-it-was-better-at-facebook-facebook-agrees.

Goodin, R.E. (2008). *Innovating democracy: Democratic theory and practice after the deliberative turn.* Oxford University Press.

Goodman, E., et al. (2017). *The new political campaigning.* Media Policy Brief 19. London: Media Policy Project, London School of Economics and Political Science.

Gorton, W.A. (2016). Manipulating citizens: How political campaigns' use of behavioral social science harms democracy. *New Political Science*, *38*(1), 61–80.

Grassegger, H., and Krogerus, M. (2017, January 28). The data that turned the world upside down. *Motherboard*. https://www.vice.com/en/article/mg9vvn/how-our-likes-helped-trump-win.

Hegazy, I.M. (2019). The effect of political neuromarketing 2.0 on election outcomes. The case of Trump's presidential campaign 2016. *Review of Economics and Political Science*, *6*(3), 235–51.

Hegelich, S., and Medina Serrano, J.C. (2019). *Microtargeting in Germany in the 2019 European elections.* https://www.medienanstaltnrw.de/fileadmin/user_upload/lfmnrw/Service/Pressemitteilungen/Dokumente/2019/Studie_Microtargeting_Germany2019EuropeanElection_Hegelich_1_.pdf.

Kalla, J.L., and Broockman, D.E. (2018). The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. *American Political Science Review*, *112*(1), 148–66.

Kefford, G. et al. (2022). Data-driven campaigning and democratic disruption: Evidence from six advanced democracies. *Party Politics, 29*(1), 448–62.

Kinkead, D., and Douglas, D.M. (2020). The network and the demos: Big data and the epistemic justifications of democracy. In K. Macnish and J. Galliott (Eds.), *Big data and democracy* (119–33). Edinburgh University Press.

Habermas, J. (2023). *A New structural transformation of the public sphere and deliberative politics.* Polity Press.

Landemore, H. (2020). *Open democracy: Reinventing popular rule for the twenty-first century.* Princeton University Press.

Lees-Marshment, J. (2009). *Global political marketing.* Routledge.

Maarek, P. (2011). *Campaign communication and political marketing.* Blackwell.

Macnish, K., and Galliott, J. (2020). *Big data and democracy.* Edinburgh University Press.

Manin, B. (1997). *The principles of representative government.* Cambridge University Press.

Mounk, Y. (2019). *The people vs democracy: Why our freedom is in danger and how to save it*. Harvard University Press.

Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin.

Przeworski, A. (2010). *Democracy and the limits of self-government.* Cambridge University Press.

Rawls, J. (1993). *Political liberalism.* Columbia University Press.

Stelzer, H., and Veljanova, H. (2020). Developing and ethical compass for big data. In K. Macnish and J. Galliott (Eds.), *Big data and democracy* (231–46). Edinburgh University Press.

Tappin, B.M., Wittenberg, C., Hewitt, L.B., and Rand, D.G. (2023). Quantifying the potential persuasive returns to political microtargeting. *Proceedings of the National of Academy of Sciences, 120*(25), 1–10.

Theviot, A. (2019). *Big data électoral. Dis-moi qui tu es, je te dirai pour qui voter.* Le Bord de l'eau.

Wooley, S., and Howard, P. (2017). Computational propaganda worldwide: Executive summary. In *Computational Propaganda Research Project,* Working Paper, 2017(11).

Zuiderveen Borgesius, F., Möller, J., Kruikemeier, S., Ó Fathaigh, R., Irion, K., Dobber, T., and de Vreese, C.H. (2018). Online political microtargeting: Promises and threats for democracy. *Utrecht Law Review*, *14*(1), 82–96.

2

Big Data
A Collective Resource in a Connected World

Pierre Trudel

Introduction

All connected objects produce data. Current privacy laws focus on the relationship between information and the individual. They are not designed to ensure that data generated by an entire population in the connected world is used responsibly. The information obtained through the "free and informed consent" of individuals is nevertheless used to generate value in the advertising market, targeted marketing, and, soon, in the prediction supported by artificial intelligence of our movements and our health condition.

As François Blais points out in Chapter 1, the foundations of democratic processes are at risk of being undermined by the use of unfair processes carried out on big data. This is a good reason to question the status of the resources mobilized to support activities that are risky for the integrity of political communication.

In this chapter, I argue that the underlying approaches of data protection laws fail to cope with the challenges posed by the advent of a context in which data produced by collectivities constitute a factor of wealth production. Hence the need to rethink the status of big data to recognize that it is a resource that emanates from the community and as such should be treated as a collective resource. Recognition of the collective status of big data as a resource provides a rationale for legislative measures intended to protect or restore data sovereignty in democratic states.

In the world of ambient intelligence and monitoring, even the most insignificant objects have the ability to connect to the network (Benyekhlef, Paquette-Bélanger, and Porcin, 2013). Digital technologies make it possible to gather, transmit, and analyze the many traces produced by the movements of all those who spend time in the connected world. Given the general trend, the reflex is to despair that we are coming to live in a "surveillance society," where the data produced by everyone's movements are captured and monetized, even at the cost of the most fundamental freedoms. Yet, from another perspective, we can imagine a legal framework for data by seeing them as a value-generating resource. Thanks to the compilation of massive data generated by portable devices, Google or Waze produce real-time information on road traffic jams. Artificial intelligence can similarly identify trends in the health status of entire populations. Whether in transport by autonomous vehicles or in leisure, activities based on the valuation of massive data in a world characterized by "ambient intelligence" are multiplying. But we too often forget that data is a resource generated by the community, which has the right to demand that it be used in a manner compatible with the individual and collective rights of all citizens.

Data are pieces of information produced by formalizing traces or signs. Data are resources that affect individuals' interests but, at the same time, because of the many ways they can be used to benefit the community, they are also a collective resource. Data are one of the primary raw materials of value-generating activities in the connected world. So far, this essential resource has been cornered for a derisory price by the major internet players.

Since is clear that data are the key resource for accumulating capital, we need to examine the characteristics of that resource, and it becomes essential to ask how we should design democratic regulation of technologies associated with real significant potential for surveillance and other freedom-destroying practices.

In a "surveillance capitalism" society (Zuboff, 2019a), the question of the legal status of resources used to create value is central. It grips much of the debate around the distribution of wealth. The way big data is used to generate value also has crucial issues for the viability of democratic deliberative processes. For example, is it compatible with democratic imperatives to tolerate that practices of valorization of massive data can generate powerful incentives to allow false information or harassing remarks to circulate online?

This new awareness of the value of data used massively in a growing number of functions related to artificial intelligence and connected objects seems rather late in coming. Why have legal analyses and ethical discussions obscured to such an extent the fact that this valuable information resource has been made available to companies who profit from it without really paying anything in exchange?

In Europe and in Canada, personal information protection laws still essentially reflect a conception of late-twentieth-century computer technology. We continue to think of information about a person as an essentially individual concern. Those working on the protection of personal information focus blindly on the idea that personal information is somehow the "property" of the individuals who generate it. This fiction leads them to insist on maintaining laws that require individuals to "consent" to the gathering and use of their personal information. Some even go so far as to dream about a world where individuals could "sell" their information to those wishing to use it. So long as the person consented, there would be no problem! There might be a requirement to report on the lawfulness of data processing, as in the General Data Protection Regulation (GDPR; see de Terwangne, 2018), but that legal framework remains based essentially on a vision of data as pieces of information attached to individuals who would make, a priori, the decision of whether or not to allow the information to be used.

The model embedded in personal information protection laws is not designed to regulate the increasingly complex processes based on the use of aggregated personal information. For example, we are dismayed that the accumulation of information about an individual can lead to discriminatory decisions. However, we very carefully tiptoe around the elephant in the room; namely, how to adjust legislation to provide people with effective protection against the kinds of discrimination that can result from data-based analyses. Personal information protection laws protect individuals against uses of their personal information, but they do not protect against profiling and other discriminatory processes. They only give individuals a theoretical right to consent, or not consent, to the collection and use of information that concerns them. Yet, how could it be legal to "consent" to information processing that could have effects prohibited by antidiscrimination laws?

Based on consent and, to a certain extent, on the obligation to show the legitimacy of the processing, the model embedded in personal information protection laws results in almost all the information circulating on the internet being collected with the "consent" of the individuals who have clicked on the ritual "I agree" button every time they go online. That is all the laws require. Additional provisions have been tacked on to some personal information protection laws to impose stricter conditions on those who deploy processes at risk of revealing something private about individuals or designed to make decisions respecting them. However, if laws truly protected both privacy and other basic rights, they would place much stricter conditions on all the mechanisms that use personal information, especially aggregated data, to create value.

A Collective Resource

When data are aggregated and analyzed, especially for the purpose of identifying global trends, taking an individualistic approach to personal information renders us powerless to implement a legal framework for protecting human rights. This is because massive data processing is based on logic different from that underlying today's personal information protection laws. The information we provide when we make queries in a search engine is of course personal information a priori. However, when that information is merged with other information to deduce that, in a specific area at a specific time, there may be an epidemic, it becomes difficult to consider that we are still in the world of personal information.

When information is detached from each individual and used to measure mass phenomena, it is absurd to consider it as personal information. It becomes a collective resource that must be regulated like a collective resource, not as a sum of information on individuals. This means that the legal status of the information becomes a matter of concern. Is it the case that, a priori, information comprises a set of objects that can be observed, compared, and compiled and that they are subject only to the rights that individuals may claim with respect to them? If information is aggregated and analytical processes are applied to calculate preferences and measure attention in network-based environments, it takes on the aspect of raw material. It is a fundamental resource for carrying out an activity; for example, to enable targeted advertising or personalized recommendation systems.

Many aspects of information involve collective interests. Information is one of the collective resources that are now essential to the functioning of contemporary connected environments. As a collective resource, information is a source of value creation, and, as such, it can be a source of wealth. The use of big data to feed analysis processes aimed at identifying trends in behaviour likely to have an impact on the community carries undeniable public policy issues. For example, the use of data from keywords introduced into a search engine can be used to detect an epidemic. Similarly, the use of big data to feed recommendation algorithms to users of online platforms for broadcasting TV shows or musical works involves cultural policy issues that go beyond mere commercial interests.

By nature, data are elements of information that flow or are taken from the observation of an individual or group of individuals. They are therefore a resource that is a priori available for anyone with the appropriate infrastructure to collect them. Insofar as data is collected in the context of activities taking place on their territory or used with regard to persons under their jurisdiction, states have a legitimate interest in regulating its use.

We can take for granted that aggregate data are a resource that can be considered of interest to the community. This provides a new rational foundation for state intervention. Access to data is a source of wealth. The pure and simple refusal to consider data sharing is an irresponsible posture. The sharing of massive data is an intrinsic characteristic of the connected society. For example, the proposals to allow, subject to strict conditions, the use of medical data for purposes associated with biomedical research illustrates the need to endow such a resource with a less naive status than that provided for by the laws on the protection of personal information. Unless we resolve to submit to the standards imposed by multinational technology companies, we need laws imposing real obligations on those who share data.

Understanding the legal framework of connected environments requires identifying the extent to which laws regulating the collection, circulation, and processing of information apply. However, more than that is needed. The legal framework has to be analyzed in a way that takes into consideration the stakes and risks created by connection in a network. We also have to specify the legal status of the resources used in producing value in connected environments.

To envisage the normative framework of connected environments in the context of the risks the technology seems to create, we have to pay attention to the characteristics and operating logic of such environments. Doing so makes it possible to identify what is at stake and give ourselves the means to pinpoint the characteristic features of the legal frameworks concerned. Networks create and organize the space within which information can circulate. The legal framework regulating information management defines the rights and responsibilities of the individuals and service providers involved in the various operations in which information is produced and used.

On the legal level, the result is a regulated space. On the technological level, it is a space governed by norms. The legal regime makes it possible to situate the personal information protection that must be ensured where the real stakes lie. The legal framework also makes it possible to specify the respective responsibilities of all who find themselves in control of information in network-based spaces (Trudel, 2009).

In the connected world, just like in the offline universe, public policy must be based on systems of justification. Not only are there risks that criminal use will be made of the mechanisms by which information circulates online, but profitable online circulation of much content depends on the processing of aggregate data. However, big data is a collective resource. Thus, several studies report systemic biases that can be perpetuated and accentuated by processes that fuel big data. Inadequate supervision of the use of big data can contribute to aggravating discriminatory biases that undermine respect for fundamental rights

(Chouldechova, 2017). The community has an interest in ensuring that the use and commodification of such resources does not result in harm to democratic values and human rights.

In light of this, we have to ask what is the status of this resource? Massive processing of the information that everyone produces is a major source of wealth creation. Identifying the status of that resource is therefore crucial. The current legal framework for protecting personal information postulates that personal information matters belong to the realm of privacy protection. However, that legal framework is not designed to regulate the production and sharing of value produced by information. To remedy this, we need to identify both the foundations for and functioning parameters of a status for aggregate data seen as a resource in connected environments.

Personal Information Law: The Late Twentieth Century Paradigm

We are used to considering the regulation and legal nature of personal information as belonging to the realm of privacy protection. We automatically look to the authorities charged with applying personal information protection laws and expect them to monitor and regulate what happens to the personal information generated by online activities. This could be justified when information concerning individuals was used mainly to make decisions concerning specific individuals. However, once information become a value-generating resource, it becomes difficult to see it as uniquely concerning individuals.

The right of personal information protection developed in the last quarter of the twentieth century, in tension between two broad points of view (Poullet, 1991). One automatically linked the right of protection of personal information to human dignity and individuals' right to control information concerning themselves or affecting their privacy. According to that point of view, law's role was to give individuals prerogatives in order to protect their dignity as humans, and it would be inconceivable for that type of prerogative to become an object of trade.

The other point of view advocated giving subjects a right with respect to personal information that could be traded on the market. It would not be a pure property right but a set of prerogatives that would include incentives to consider the sharing of the benefits that can be extracted from information about our personal lives.

At the time personal information protection laws were established in most democratic states, the idea that personal information could be an object of trade was rejected. Such information's connection with human dignity made it inconceivable that it could be commodified. This is why the possibility of collecting and using information had to be subordinated to obtaining each individual's informed consent.

Individual Consent

At the time, there was insistence on the need to protect personal information from becoming something that could be traded on the market infinitely. This is why such information had to be seen as a component of human dignity. Law had to protect the individual's freedom to control information about them. This is why it seemed unacceptable that this type of information could be subject to property rights likely to be traded on the market. The legal community's usual reflexes were engaged to design legal instruments to protect personal information. In the quest for a legal vector that would give individuals the ability to control their personal information, we automatically resorted to a centuries-old legal mechanism for transactions concerning extra-patrimonial rights. The mechanism is based on the subject's consent, and it is the device that personal information protection laws use.

The extra-patrimonial model based on the individual's "free and informed" consent has led to the development, consolidation, and sclerosis of laws on personal information (and data) protection. Those laws are all structured around the requirement that the individual concerned must give "free and informed" consent. To use information obtained from or captured about individuals, their consent must be obtained. This legal operation based on the "meeting of minds" is how individuals' right to control information about them is embodied.

In such a system, primary importance is placed on ensuring that the individual's consent has truly been obtained and that they have not been "tricked." This leads to a legal framework that functions so as to preserve or restore fair play in an individualistic context. It is as if there were a pact between the individual and the entity that collects and “processes” the information about them. Marked by the postulate according to which a piece of information may be held for a specific purpose, along with other rules designed to ensure that the individual keeps control over information concerning them, the legal framework for personal information protection developed in this way, always further formalizing the requirements supposed to protect the individual’s right to control information about themselves.

The Trivialization of Consent

The spread of the internet and the connected world has exacerbated the need for rules controlling the circulation of personal information in open networks that lack the forms of protection that are available, by default, in traditional information processing environments. In a sort of panic, we have simply thrown overboard the distinction between information concerning one’s private life and information concerning one’s interactions with others, with one’s peers. However, it turns out that once aggregated, the personal information that becomes big data concerns the community much more than each individual taken in isolation.

The result is that we take rules and reasonings that were designed to limit the use and circulation of information pertaining to privacy and apply them to information circulating legitimately in public. In the name of protecting personal privacy, we have even found it legitimate to erase (or force the forgetting of) information on an individual’s participation in collective life. Recognizing a right to erasure of personal information when maintaining it in online spaces may cause individual harm. Over the decades, personal information protection laws have spread their reach, requiring the subject’s consent for any “processing” of personal information (Gautrais and Trudel, 2010). An independent category of information object has even emerged: personal data. In some legal systems, particularly in Europe, this has replaced a number of other concepts designed to explain the stakes and balances that must be taken into account in the circulation and use of information about individuals.

The notion of “personal data” flows from a conceptual slide motivated largely by concern to protect the subject’s ability to control information about their characteristics and identity. Designed initially to maintain equilibrium within centralized bureaucratic organizations, the notion has been put to use to provide a framework for the flows of information circulating on the internet, but this has been at the price of conceptual simplification. The notion of "personal data" means we must discard the flexibility of the notion of privacy. It ignores relationships with the information context in which the information circulates. It eliminates the many-sided nature that is necessarily an aspect of information about persons. While privacy issues are usually analyzed by the courts taking into account the context to determine whether information about an individual can circulate publicly, laws on the protection of personal information rule that information relating to an individual is necessarily confidential.

This simplification becomes material in the complete disappearance of the formerly obvious distinction between private information and information that concerns other people, those close to us and the public. This simplification has led to many slides, including, notably, simply tacking the legal framework initially designed to protect information concerning the private sphere and intimate life onto public information. The persistent description of personal information as an extension of the individual has complicated the development of legal approaches able to deal with what has become the information resource in the connected world and with the many issues that it raises in government administration and in the functioning of connected activities (Trudel, 2008).

The consent-based paradigm does not prevent us from lamenting the fact that, in many services associated with the "information society," the individual is now the commodity being sold. Social networks and search engines offer services for free but require "consent" in counterpart so as to transform an extra-patrimonial object into a value-generating resource. Given this paradox, "surveillance society" becomes the target of much condemnation. But is taking a fatalistic view the only answer? Is it not time to recalculate our legal course so as to return to a balanced path? By overinvesting in a legal regime that does not take into consideration the relationships between information and individuals, the legal communities in most democratic countries (especially in those that have made personal information protection their stock in trade) have delayed the development of a legal framework that would have ensured a better distribution of the value generated by data.

Expropriation of the Value of Personal Data

By clouding the fact that information and data generate value, personal information protection law has prevented itself from regulating the processes by which we generate value in connected environments. This is a fact, for we cannot overlook the reality that, in practice, today's legal framework allows personal information to be considered a negotiable commodity. Data are bought and sold every day, no matter how long we drone on repeating that they are noncommercial resources.

Every day, information obtained thanks to individuals' presumably free and informed consent is a good that is traded, and this is despite the formal legal framework that is supposed to ensure it remains extra-patrimonial. Formal appearances are saved. However, the reality is very far from the premises of the formalist framework that is supposed to protect individuals by protecting their personal information.

The commodification of data is a feature central to the functioning of digital society. It is made possible by the spread and trivialization of individuals' "consent." The legal framework that is supposed to protect individuals from the commodification of personal information that is a component of human dignity has proven to be a powerful vector for the expropriation of the value of data. Each individual's consent it required, but that is all that is required to extract value from the data deposits created by the actions and things done by all those who live on the network and in the connected world.

The value of that precious resource is appropriated and nothing is given in return, aside from compliance with a few formalities. Worse, this kind of legal framework leaves the operations by which value is extracted from big data without any meaningful supervision. The lack of a significant legal framework allows "accidents" to happen and balance to be lost (as evidenced by scandals such as Cambridge Analytica), and then we are left picking up the pieces.

Of course, we can always have fun "controlling" our personal data individually, changing our default settings, even exiting Facebook or pretending to have had enough of the connected world. We can say we will stop using Google and Instagram immediately! But this is all artifice that naively supposes it is all just an individual affair, only a "contract" between cybernauts and a company that swears it has our data at heart. These paradoxes suggest we should consider the data produced by the individuals who act in the connected world from a collective point of view. This means looking at the legal operations that are characteristic of surveillance society.

The "Surveillance Society" Paradigm

The notion of surveillance capitalism was developed in 2014 by American economist Shoshana Zuboff, professor emeritus at Harvard Business School. In the book she published in 2019, *The Age of Surveillance Capitalism*, she describes the strategy deployed in less than twenty years by internet groups that were propelled by both neoliberalism and the increasing acceptability of mass surveillance following the attacks of September 11, 2001. The increased ability to quantify and analyze data and the resulting greater predictive power threaten individual freedoms and democracy.

Zuboff (2019b) describes the shift towards a new era of capitalism. In her book, she makes a connection between twentieth-century industrial capitalism in Ford's car factories and the form of capitalism invented by Google at the turn of the 2000s. She argues:

> The digital industry prospers thanks to a principle that is almost childish: extract personal data and sell advertisers predictions about users' behavior. However, for profits to grow, the forecast has to turn into certainty. For that, predicting alone is no longer enough: it is now a question of changing human behavior on a grand scale. (10)

Algorithmic processes based on artificial intelligence technologies have considerable capacity to collect, compile, and analyze data from many different aspects of each individual's life and to infer, or even anticipate, an individual's behaviour. Given this potential, people often express their resignation or outrage that we are living in a surveillance society. Some call for individual action to reject technology and others advance conspiracy theories that demonize technological objects (Laugée, 2019).

By continuing to consider the information that every one of us produces as an individual resource, a resource that individuals are responsible for protecting or for "managing" by giving informed consent to their use, we encourage the practice by which the value of information is expropriated. Individuals are dispossessed of the value of their information, and the value of that resource for the community is simply appropriated without compensation by the companies that control the technology for capturing and processing it.

Understanding the functioning of surveillance capitalism and recognizing the crucial importance of data and information in the wealth creation process is a necessary condition for identifying the controls that need to be imposed by law. However, it is not sufficient. We need to go further and identify how to ensure the rights are distributed among the various stakeholders concerned by the information, the resource, that is now fundamental to the creation of wealth.

We need to go beyond fatalistic discourse and identify ways and means to regulate such companies. We need to realize that, for a large part, the reason surveillance capitalism has a clear path is because of the belief that it is impossible to regulate the processes by which monopolies expropriate the value of data without any real compensation. We urgently need to re-establish state capacity to regulate the activities of surveillance capitalism society. For this, we have to set up legal mechanisms that will determine the rights and obligations of persons, individuals, companies, and leaders in the production of value from data. This supposes identifying the resources that are fundamental to the creation of value and distributing the rights and obligations with respect to those resources appropriately. However, at this time, we are lacking analysis regarding the status of information as a value-generating resource in a surveillance capitalism society.

Aggregated Data

In the connected world, information is increasingly seen as a resource generated by the community. It should therefore be regulated in the interest of all. Marie-Anne Frison-Roche (2016) suggests thinking about the world from the starting point of the notion of "data." Pointing out the poor fit of the *summa divisio* between economic regulation and public freedoms, she asks to what extent law, understood as a legal system, is calibrated appropriately to deal with a world that includes ambient intelligence.

Why postulate that the value they make it possible to generate is necessarily private property? Should not the personal data produced by objects, movements, and other situations involving the actions and things humans and objects do in connected spaces be seen as having the characteristics of a collective resource? If so, then their commodification could be used to meet the community's needs. For example, the big data produced by the set of actions and things done in living environments is a resource generated by all of the interacting beings, as well as the objects that are under their direct or indirect control. Similarly, targeted advertising as well as other decision-making processes based on algorithms – in short, all the processes using data produced by the actions of the community to create value – should be regulated in order to guarantee that the extraction of value from the data is done in accordance with the general interest.

Of course, such pieces of information can, when used to obtain a way of identifying an individual, become subject to privacy protection requirements. However, when they are raw materials for value creation (for example, when they are used to support the broadcast of public programming or to generate playlists offered by Spotify), they constitute a resource in which the community has an interest. The community's interest is generally much greater than that which could be claimed by an individual. Like radio waves at the turn of the twentieth century, data are starting to look like a resource that belongs to the community and on which it is legitimate to impose, on behalf of the community, fiduciary duties in the community's favour, at the very minimum.

Changes in the Legal Nature of Data

Personal information protection protects information that is closely linked or connected to a given person. However, when information is aggregated into big data, its nature changes. The notion of personal information concerns specific objects, but as data become big, they no longer have a specific object. Like the air we all breathe, data are personal when they are really connected to an individual. The air that a person absorbs is part of that person so long as the precious molecules remain related to their body. When the person breathes the air out, the expelled molecules are no longer part of them; the molecules become part of the environment, a resource that is of critical concern to the community. Of course, we have to ensure that what comes from an identifiable person cannot be used to harm them, but, when aggregated, the information we all produce through our actions and the things we do, by what we are, involves crucial stakes, and those stakes are not all individual. This is why some say that data are the oil of the twenty-first century.

This point of view provides legitimate foundations for demands in favour of imposing consideration for data use. Why should access to these communal resources remain free of charge? Is it not legitimate for communities to have something to say regarding the value generated from pieces of information produced by the movements and interactions of the many individuals belonging to them? By considering big data as a collective resource, communities and states would be able to attach conditions onto the processes by which value is generated from data.

For example, processes such as targeted advertising and other algorithm-based decision processes – in short, all value-creation processes using data produced by actions and things done in the community – should be governed in a way that ensures the extraction of value from the data deposits remains consistent with the general interest.

An Example: The Regulation of the Online Audiovisual Industry and Online Platforms

In a world where audiovisual products are broadcast in online platforms, the status of data as a collective resource gives rational foundations to legislation designed to ensure the availability of original Canadian works. Original creations can create value not only using frequencies but also, and increasingly, using data to measure the attention of connected individuals. When they are aggregated, the data become a collective resource, the origin of which is observation of the actions and things done by those living in Canada.

Classically, broadcasting legislation was a mechanism for determining rights and obligations with respect to the value generated by making content available, in the form of programs. The content was used to attract attention and extract income from the fact of attracting attention. Until the end of the twentieth century, frequencies were the primary resource used to create value by broadcasting or distributing programs (Trudel and Abran, 1994). With the advent and spread of the internet, individuals' attention became the resource used to generate value. By measuring the actions and things done by the individuals living in the country, data about individuals' attention can be produced and analyzed.

The Canadian broadcasting legislation states that broadcasting undertakings form a system. Section 3(1)(b) provides that the "elements" of the system, most of which are companies, "make use of radio frequencies that are public property." Today, radio frequencies are no longer the principal public resource required for the operation and, especially, the creation of value in the context of broadcasting. The internet has created a new context, one that no longer operates like the broadcasting environments referred to in broadcasting legislation (Trudel, 2015).

Content now circulates in a connected environment. Users now control what they watch or listen to. The function of programming (curating) pieces of that content is now in the hands of each user. The idea that content is programmed according to a time schedule is being replaced by a platform system on which broadcasters and users interact and users make choices according to their preferences.

Identification of Content

In this context, what is crucial is to ensure an environment in which each individual is truly able to exercise significant choice; for example, by being able to access content from all sources, including Canadian ones. This supposes making sure by putting in place appropriate regulations that when Canadians pay to consume audiovisual products, their money is in part directed toward the production of content reflecting Canadian creativity. A system that could not guarantee the viability of the production of Canadian content could not be considered as allowing real choice. From this perspective, it is essential to guarantee the availability and the discoverability of content made by Canadian creators.

However, content's discoverability depends on data making it possible to identify, locate, and position documents in a network-based environment. To act on discoverability, the legal framework must make it possible to put in place measures, and the necessary authority, to regulate the production and use of data associated with the production and consumption of content, in other words, descriptive data and use data.

These resources are now essential elements in the operations of companies that produce and distribute information content. It seems clear that we will have to identify the extent to which audiovisual legislation should authorize regulators to establish conditions for the production, sharing, and use of data on available content. When it comes to discoverability, for content producers and commodifiers, users' attention is an essential strategic resource. Producers and broadcasters want their content to be available and, especially, visible to users. This means that they are in competition to quantify and commodify the attention of those active in the connected environment. This means that users' attention is a crucial resource for value production. And value creation is an essential condition for the funding of productions expressing Canadian creativity.

Calculating and Valuing Users' Attention

Attention generates value. In fact, when we think about it, attention has always been raw material for value-creation processes in information environments. In the past, attention was calculated in terms of newspaper and magazine circulation and radio and television audience measurements. Those figures were used to establish the value of content on the advertising market. A program's value was established according to the size of its audience. By selling advertising space and generating revenue, media companies realized and created value.

In a connected environment, data are the instrumental resource for measuring attention. Today, attention is measured by gathering and analyzing data. Data are pieces of information resulting from the observation of the actions and things done by everything and everyone who is connected to the network. The analysis of the data, using analytical processes applied to aggregated data, is a necessary condition for the creation of value.

The example of online broadcasting platforms shows how data as a resource flow in the value generation cycle characteristic of connected environments. The legal framework must be brought up to date to reflect the interests that need to be protected at each step in the data commodification cycle.

A Collective Resource and Sovereignty Issue

Data is obtained by observing and compiling the actions and things done by persons and objects connected to the network. When the data-generating observed actions and things take place within national borders, they are objects that must be protected on behalf and in the interest of the community. As such, big data becomes a common good. It is used by many, but like common goods such as air and water, it must be used in the community's interest. Owing to data's collective nature and the rights that the community can claim, their regulation is a sovereign power. Since it is produced in part by the observation of what happens in spaces within national borders, big data is by nature a collective resource and subject to collective rights. The connection to sovereignty issues nonetheless requires first that data concerning identifiable individuals be distinguished from data that have been aggregated to such an extent that it is impossible to identify any individual (without using re-identification techniques).

Pierre Bellanger (2016) explains that personal data are pieces of information that inform us, directly or indirectly, about an identified person. That definition establishes a unique personal right of the person concerned with respect to their data, a right intended, in particular, to protect their privacy. Legal and institutional reflections tend to seek to strengthen and confirm the exclusive right that each person has with respect to data concerning them. However, personal information laws postulate that personal data provide information on only a single individual. Bellanger notes that such laws "consider personal data to be granular, independent and forming an entity in itself, subject to the right of only one" (15–16).

Acknowledging the dual nature of data concerning individuals, Danièle Bourcier and Primavera de Filippi (2018) suggest investigating the individual and collective natures of the right to privacy. According to them:

> On one hand, it is a right relating to a person's private life, a profoundly individual right. On the other hand, however, the right to

> privacy has acquired a new meaning with the advent of digital technology and the Internet. (456)

Bourcier and de Filippi argue that in a hyperconnected world, the right to privacy can no longer be seen uniquely as an individual right. It must be interpreted as a collective right; in other words, as a right that is related to individuals but that must be exercised in common or in a collective manner. Julie E. Cohen (2019) draws attention to the fact that, when we analyze privacy protection stakes, we must take into account information's intrinsic mobility. She writes:

> In the networked information era, preserving effective privacy protection for the subject of privacy entails decentering them both within theoretical frameworks and in the design of privacy institutions. Developing institutions and practices for operationalizing an affordance-based approach to privacy requires a design ethos informed by careful attention to the relationship(s) between materiality and practices of self-articulation and especially to the importance of boundaries, gaps, and discontinuities for those practices. It also requires a robust conception of operational accountability that moves beyond individualized choice and consent to emphasize responsibility, respect and new modalities for effective regulatory oversight of algorithmic and data-driven processes. (22)

Data are multifaceted. They are no longer attached only to individuals. They now constitute a network. Bellanger (2016) notes that each datum remains personal, but the data are now organized into an unfragmentable whole. Personal data "cannot be isolated in practice: giving access to one's contacts, photos, agenda, email and location automatically involves the personal data of other people, over which one has no right" (16). He also notes:

> Personal data inform on other people: correlation algorithms, which are computer programs that use probability to deduce information through predictive aggregate processing of data with no direct relationship to the information inferred, mean that every personal datum

> indirectly provides information on other people. For example, the personal data of bank clients matched against their payment defaults will be used to calculate the risk that new clients will default on the basis of comparison of their behavior. Another example would be where data on colon cancer in a group of individuals correlated with their supermarket purchases makes it possible to predict an individual's risk of cancer, even if that individual has no relationship to the test group, on the basis of their sales receipts alone. (16)

In practice, in such an environment, individual control through the supposedly free and informed "consent" mechanism becomes illusory. Personal data collection will grow. For example, numerous sensors scattered everywhere and integrated into most objects record and process various data. Like others, Bellanger observes that "thought-out authorization for each collection, which is already hit or miss, is no longer possible in fact" (17).

In fact, it is unrealistic to imagine that the distribution of the value resulting from data could flow uniquely from individual consent, even supposing that such consent has real meaning. That said, we can agree from the outset that human dignity issues can be taken care of by personal information legislation. However, collective and industrial issues, such as those related to the availability of content, fall under sector-specific legislation, such as legislation on broadcasting and the audiovisual industry.

Moreover, data, whether they are personal or aggregated, are subject to the network effect. Bellanger (2016) notes that their value is proportional to the square of the number of data to which they are connected. This means that the value of a datum depends mainly on its context, which is itself a generator of additional data. Bellanger writes:

> A unit of data has no absolute value. However, the biggest data holder can constantly outbid others, in cash or services, to acquire new data because, for the biggest data holder, data are what have the most value and each new acquisition increases the value of the set already collected. By this logic, the biggest data holder will continue in this way until he has a monopoly. (16)

Therefore, the appropriation by a few companies of the information-based reconstitution of reality is a source of devastating competition asymmetry and cannot be prevented by a sum of individual rights. The result is that personal data are no longer granular but networked. They thus involve a number of collective issues. In the audiovisual world, they raise major strategic stakes.

This shows how data can be seen as a common good, as property in which there is collective interest. This provides a rational foundation for rules to regulate the ability of companies to extract value from processing big data. In other words, data require a framework in the context of the relationships between individuals and the enterprises who collect and use individuals' data. However, we also have to consider data as a collective resource when, once aggregated, they become a value-generating resource. Every state can legitimately claim a right to exercise oversight with respect to data resulting from the actions and things done by those located in whole or in part within its borders. For these reasons, data must be considered not simply as something that concerns each individual, but also as a collective resource. Data directly linked to an individual are taken care of by personal information protection laws, and big data used to generate value in companies involved in the broadcasting and distribution of audiovisual content falls under legislation that regulates telecommunications companies and the audiovisual industry.

As a report by a task force on research in science and digital technologies (Ganascia, Germain, and Kirchner, 2018) involving a number of French research institutions explained, the fact that a number of activities have shifted into networked space has changed the conditions for exercising national sovereignty. Formerly founded on controlling what happens in the state's physical space, sovereignty now depends on the ability to control what goes on in the networks. The report recalls that "sovereignty can be defined as an entity's ability to give itself its own rules or, more trivially, as the 'power to exercise power'" (10). It further observes that the conditions created by the preeminence of network-based environments forces us to wonder about "the digital sovereignty of States, organizations and individuals ... or forms of supranational sovereignty" (10).

The “sovereignty of data” is one of the most obvious issues. Individuals’ personal data must be protected, but the big data now used to create value is a resource that concerns both individuals and communities living within national boundaries. When data circulate in networks that have become largely unconcerned by territorial borders, state sovereignty is in trouble. States need to ensure that the choices flowing from their citizens’ values prevail and they have to give themselves the means to impose rules that are consistent with those values.

Conclusion

Data produced by individuals and the compilation of those data to calculate community trends are fundamental raw materials for the network-based functioning that characterizes our era. The fact that many activities function in networks makes it more complicated for states to control what is going on within their borders. In this way, an increasing amount of data escapes any form of democratic control. The conditions produced by the preeminence of network-based environments force us to ask questions about the digital sovereignty of states, organizations, and individuals or forms of supranational sovereignty. Sovereignty must now be exercised in spaces that are increasingly virtual. It is important that states give themselves the means to ensure that the choices and rules that prevail are the ones that are consistent with their values. They must act alone but at the same time innovate by giving themselves the means to act in concert with other states.

However, the establishment of relevant intervention with regard to processes that operate on the commodification of big data requires breaking down certain epistemological barriers that affect analyses concerning data. Seeing data concerning individuals as a resource that concerns the individual alone, who would be “free” to consent to their use, is unrealistic when data are aggregated. Nonetheless, the mere fact that big data is designated as a collective resource does not entail that public authorities are justified in accessing information on individuals. We need to insist on the distinctions between the designation of a resource as being collective and a state’s right to access data when those data make it possible to identify individuals.

Individuals' personal information must be protected, but aggregated data, which are now used to create value, are a resource that concerns both the individuals and the communities living within national borders. State sovereignty is threatened because the data circulating on networks are generally unimpeded by territorial boundaries. This is why it is important to recognize the sovereignty stakes raised by data. The advent of the connected world forces us to re-examine the processes by which state sovereignty is expressed and state laws are applied. In a networked world, technical configurations and the way objects are connected lead to default regulation of behaviour. Regulation by default means the requirements imposed by the technical configurations installed in the objects must be scrutinized in the framework of open processes designed to establish mechanisms able to guarantee accountability.

References

Bellanger, P. (2016). Les données personnelles : une question de souveraineté. *Le Débat*, (1), 14–25.

Benyekhlef, K., Paquette-Bélanger, E., and Porcin, A. (2013). Vie privée et surveillance ambiante: le droit canadien en chantier. *Droit et cultures. Revue internationale interdisciplinaire*, (65), 191–223.

Bourcier, D., and de Filippi, P. (2018). Vers un droit collectif sur les données de santé. *Revue de droit sanitaire et social*, *2018*(3), 444.

Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data*, *5*(2), 153–63.

Cohen, J.E. (2019). Turning privacy inside out. *Theoretical inquiries in law*, *20*(1), 1–31.

de Terwangne, C. (2018). Les principes relatifs au traitement des données à caractère personnel et à sa licéité. In C. de Terwangne and K. Rosier (Eds.), *Le règlement général sur la protection des données (RGPD/GDPR): analyse approfondie* (87–142). Larcier.

Frison-Roche, M-A. (2016). Penser le monde à partir de la notion de "donnée." In M.-A. Frison-Roche (Ed.), *Internet, espace d'interrégulation* (7–16). Dalloz.

Ganascia, J.-G., Germain, E., and Kirchner, C. (2018). *La souveraineté à l'ère du numérique : rester maîtres de nos choix et de nos valeurs*. Commission d'Ethique sur la Recherche en sciences et technologies du Numérique d'Allistene (CERNA). https://www.allistene.fr/publication-de-la-cerna-sur-la-souverainete-a-lere-du-numerique/.

Gautrais, V., and Trudel, P. (2010). *Circulation des renseignements personnels et Web 2.0*. Éditions Thémis.

Laugée, F. (2019). Capitalisme de surveillance. *Revue européenne des médias et du numérique*. https://la-rem.eu/2019/07/capitalisme-de-surveillance/.

Poullet, Y. (1991). Le fondement du droit de la protection des données nominatives : Propriétés ou libertés. In E. Mackaay (Ed.), *Nouvelles technologies et propriété* (175–205). Litec, Éditions Thémis.

Trudel, P. (2008). Hypothèses sur l'évolution des concepts du droit de la protection des données personnelles dans l'État en reseau. In M.V. Pérez Asinari and P. Palazzi (Eds.), *Défis du droit de la protection de la vie privée: Perspectives du droit européen et nord-américain* (531–58). Bruylant.

Trudel, P. (2009). Privacy protection on the internet: Risk management and networked normativity. In S. Gutwirth, Y. Poullet, C. de Terwagne, and S. Nouwt. (Eds.) *Reinventing data protection?* Springer.

Trudel, P. (2015). Les mutations internationales de la régulation de l'audiovisuel. In S. Regourd and L. Calandri (Eds.), *La régulation de la communication audiovisuelle : enjeux et prospectives* (123–39). L.G.D.J., Institut universitaire Varenne.

Trudel, P., and Abran, F. (1994). Le caractère public des fréquences comme limite à la liberté d'expression. *Media and Communications Law Review*, *4*, 219–58.

Zuboff, S. (2019a). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.

Zuboff, S. (2019b). Un capitalisme de surveillance. *Le Monde Diplomatique*. https://www.monde-diplomatique.fr/2019/01/ZUBOFF/59443.

The Democratic Specifications
Preserving Fundamental Rights and Freedoms in Human Societies in the Digital Era

3

François Pellegrini

Introduction

The computerization of administrations in the 1970s quickly gave rise to many fears about the infringement of freedoms made possible by the processing power of computers. In France, the SAFARI project, aimed at interconnecting all the administration's files around a unique identifier for individuals assigned at birth, triggered a state scandal in 1974, fuelled by the memory of the data collection carried out during World War II (Black, 2001). Popular emotion led the legislator to adopt a law in 1978 aimed at protecting people against the abuse of personal data processing, known as the Informatics and Liberties law. Article 1 of this law, still in force, sets out in a visionary way the principles to be guaranteed: "Informatics must be at the service of every citizen. Its development must take place within the framework of international cooperation. It must not infringe on human identity, human rights, privacy, or individual or public liberties."

However, more than half a century later, the lessons of history have not been learned. The productivity gains induced by digital tools have reinforced the prevalence of "technicist thinking" (Ellul, 1964), according to which the functioning of human societies must be analyzed as a set of organizational and technical problems to be solved. Its current form is "technological solutionism" (Morozov, 2013), a doctrine of computerized management of human societies. The massive collection of personal data by private actors, as part of their business practices, has paved the way for a "surveillance capitalism" (Zuboff, 2019), aimed at predicting and influencing people's behaviour (Rouvroy, 2014).

The human experience of users of digital platforms constitutes the raw material and source of surplus of this informational capitalism, which is methodically extracted (Pellegrini and Verdon, 2022). Both the data collected and the technologies mobilized are of interest to the public power for its own missions, sparking a collaborative dynamic between private and public interests for the control of populations (Mattelart and Vitalis, 2014). The techno-solutionist narrative, relying on the promise of the comfort of the "social user experience," but sometimes also resorting to the spur of fear vis-à-vis surrounding threats, encourages populations to consent to the use of increasingly intrusive devices (see Lyon, this volume, for a discussion on the relevance of the foundational principle of the social contract in the context of the COVID-19 pandemic). These devices are shaping the contours of a surveillance society, which, on the very pretext of preserving them, is eroding the principles that underpin democratic regimes (Chavalarias, 2022).

Faced with this risk, it is necessary to define "democratic specifications." In the industry, a specification is a contractual document defining a set of principles and rules to be respected by the successful bidder. In the information technology field, this document, renamed "special technical specifications," sets out the properties expected of a computer system so it fulfills its function without deviations. The purpose of this chapter is to apply this approach to the functioning of human societies in the digital era, when information technology provides states and certain private actors with surveillance and control capacities unprecedented in the history of humanity. It is a matter of proposing a set of rules aimed at preserving the democratic character of advanced societies, as well as at protecting populations in times of unrest, when the democratic rule of law is threatened or has disappeared. Each of the rules proposed in the following sections will be justified by its necessity with regard to past history and possible futures, taking as an example the recent evolution of French law.

Life and Death of Democratic Societies

Numerous computerized systems are proposed and deployed with the intention of combating threats to democratic regimes. In the face of the intensity of these threats, exceptional measures would be necessary, even if they cut back, in a way that is presented as temporary, on the fundamental rights of individuals. However, this way of thinking is deleterious for several reasons.

First, in the short term, it leads democratic societies to become accustomed to surveillance tools. These tools are promoted by two main categories of actors: law enforcement agencies, which, motivated by a legitimate desire to apprehend the guilty, regularly solicit the legislator in order to benefit from new technical means; and companies in the surveillance industry in search of new markets (Pellegrini and Vitalis, 2017). The increase in surveillance is not limited to technological devices but is also reflected in the evolution of the law. In France, for example, identity checks, which were originally impossible outside a judicial police investigation, are now widely authorized (Desprez, 2010). However, each provision that reduces public liberties has a ratchet effect, as governments are generally reluctant to repeal a liberticidal measure, even one that is recognized as ineffective, in the face of public pressure.

Second, in the long term, this way of thinking poses a major risk of undermining the democratic character of societies, whose governments increasingly resemble those against which they claim to defend themselves (Harari, 2016). Putting in place, in the name of preserving democracy, tools of a totalitarian nature constitutes a suicidal trajectory. Democratic societies are not eternal, as European history has amply demonstrated over the last two centuries. To forge tools of excessive power that can be misused for the purpose of controlling the population on the sole pretext that those who use them will always be respectful of democratic values is wishful thinking. This deleterious trajectory concerns not only the digital sector. Thus, equipping law enforcement agencies with equipment and doctrines from the military sector leads to a logic of escalation and confrontation with the population that undermines fundamental freedoms (such as the right to demonstrate peacefully) and the social contract. When force is no longer used legitimately and proportionately by those to whom the monopoly of its use has been entrusted by the people, the latter are more and more incited to challenge it. One might reply that a tyrannical government would not hesitate to forge such tools quickly. However, by the time they are in place, the elements of the population who wish to resist oppression would have time to organize. If these tools were immediately available against the population, there would be no time to adapt.

The wisdom of a democratic system lies in the renunciation of the postulate of its perpetuity (also called the "democratic postulate"). More than that, a democracy must always anticipate its own demise to organize the protection of people beyond it, with the aim of promoting its resurrection. These are proactive measures of resistance to oppression,[1] a fundamental right enshrined in Article 2 of the Declaration of the Rights of Man and of the Citizen of 26 August 1789, written at the dawn of the French Revolution.

To support the introduction of increasingly intrusive surveillance technologies, the slogan "security is the first of freedoms" is often repeated. However, as François Sureau (2022) has so masterfully summarized, this is a "stupid formula." Indeed, democratic regimes are not based on an illusory "right to security" that is impossible to guarantee but on the "right to *personal* security"; that is, the fact of not being subjected to the arbitrariness of a tyrannical power.[2] The discourse of fear, systematically mobilized to justify the use of increased surveillance, amounts to making people believe they would be happier under an authoritarian regime. These are manipulative methods that should only be used by tyrannical regimes and lead to people having neither freedom nor security, in the absence of personal security.

> *Rule #1.1: Renounce the democratic postulate.*
> *Rule #1.2: Do not resort to the tactics and means of authoritarian systems.*

Mass Surveillance

The continuous decrease in the cost of computer equipment, access to high-speed digital networks, and AI developments facilitate the deployment of electronic systems for capturing and processing data of all kinds (see Trudel, this volume, for a discussion on the potential collective use of such aggregated individual data), enabling the surveillance of populations, by nature or by destination. The processed data can be distinguished in two categories: flow data and stored data.

Flow data are those from real-time surveillance devices, such as camera systems placed in public space. Their massification leads to an informational deluge for human operators, who find themselves subject to a cognitive overload that is ultimately detrimental to the effectiveness of these devices (Mucchielli, 2018). This is why automatic analysis systems are now proposed to alert operators to the occurrence of events supposedly "of interest." This algorithmic assistance processing, added to the capture devices thus "augmented,"[3] are supposed to allow an operator to monitor more areas at a time, with only the information requiring their attention to be presented to them. This processing, which algorithmically encodes the social norm, thus leads to considering as suspicious everything that is "abnormal."

The installation of surveillance devices in public space generally leads to a subsequent increase in their number. In the case of video "protection," the fact that these devices are not significantly effective in practice (Gormand et al., 2021; Mucchielli, 2018) is interpreted as the need to deploy more of them to achieve the desired effect, according to the argument that their ineffectiveness would be a phenomenon of a quantitative rather than qualitative nature. Prescribers find themselves trapped in an engagement strategy so as not to disavow their initial decision.

Stored data, on the other hand, has three main origins: technical traces, data voluntarily shared by individuals, and data produced by the public space surveillance tools already described (by recording flow data). The desire of law enforcement authorities to be able to legally retain and exploit this data for surveillance purposes is at the origin of many legal provisions.

Technical traces represent a considerable deposit of highly informative data on the behaviour of individuals and is therefore both extremely intrusive and much prized by law enforcement. However, mass surveillance produces profoundly deleterious effects, as the Court of Justice of the European Union (CJEU) noted in a 2014 ruling that has become famous:[4] "The fact that data are retained and subsequently used without the subscriber or registered user being informed is likely to generate in the minds of the persons concerned the feeling that their private lives are the subject of constant surveillance."[5] In its successive rulings (Tracol, 2023), the CJEU has outlined a framework for reconciling the protection of the rights of individuals with the interests of the public authorities (Tracol, 2021): data may only be collected and used in a generalized and indiscriminate manner for the sole purpose of guaranteeing national security;[6] on the other hand, for the fight against serious crime, data may only be collected and used in a targeted manner and under the control of an independent authority.

States are not lacking in imagination either when it comes to exploiting the data voluntarily shared by individuals. For example, the French General Directorate of Public Finances has been granted, within the framework of a three-year experiment, the authorization to carry out "computerized and automated processing allowing the collection and exploitation of data made public on the websites of online platform operators."[7] Once known to the public, these systems lead to the reduction of freely accessible information on the internet, by the self-censorship of internet users and restriction of access to the information they publish, indirectly infringing on the freedom of expression. The relevance of the processing implemented by the administration on the data it holds also raises questions. In the Netherlands, the implementation by the tax authorities of a biased algorithmic fraud targeting process led to the so-called Toeslagenaffaire scandal.

The solutionist postulate leads to a preference for technological surveillance tools over human means. The alleged ability of machines to detect "weak signals" within large sets of behavioural data that would escape human analysts encourages the mass collection and centralization of the information collected. This is why the general architecture of surveillance systems is moving towards an increasingly centralized model in which information is concentrated in fewer and fewer human hands. This centralization poses, in the long term, a major democratic problem. Indeed, the fight against tyranny, recognized by the right of revolution against it, is one of the founding principles of modern democratic regimes. However, for this principle to be implemented, a sufficient number of people must be able to unite and weigh in against an oppressive power. This capacity is severely undermined by the means of mass surveillance, which make it possible to identify almost instantaneously the social networks and the usual interlocutors of a person as soon as they use digital tools.

In recent history, resistance to oppression has been built by many individuals who, in their positions, have sabotaged the information gathering organized by their adversaries.[8] This individual capacity for resistance is attacked when the tools of surveillance come under the control of a small number of carefully selected and easily monitored individuals. A single person must not be able to monitor too many, otherwise it will become impossible to escape the dictatorship of an ever smaller minority. The tools and processes of mass surveillance must therefore not be able to "scale up."[9] The deliberate technical and organizational inefficiency of these devices is a desirable feature of a democratic system because it increases the number of people who have to collaborate in their proper functioning. Moreover, the use of these tools must be strictly supervised and monitored by supervisory bodies vested with extensive powers.[10]

Rule #2.1: Minimize the use of mass surveillance.
Rule #2.2: Strictly regulate and supervise the use of mass surveillance.
Rule #2.3: Do not implement scalable surveillance systems.

Regulation of Encryption Means

Encryption of communications allows interlocutors to preserve the confidentiality of their exchanges via untrusted communication channels (Singh, 1999). The spread of digital services and electronic commerce, which require reasonably secure data communications, has led to the gradual liberalization of digital encryption tools. The implementation of passwords and encryption of digital media are now recognized as good practices for both individuals and companies, so that the loss or theft of devices cannot lead to leaks of sensitive data.[11] Device manufacturers highlight these features in their sales arguments as proof of the attention they pay to the protection of their customers' privacy, to the point that personal data protection authorities are now likely to consider it negligence for a data controller not to use them.[12]

In contrast to this broad move, voices have raised to oppose the generalization of encryption techniques. Criticism comes mainly from law enforcement agencies and judges who are worried about neutering the means of interception used in the fight against terrorism and serious crime. These fears have encouraged legislators in some countries to demand the implementation of "backdoors" in encryption means, so that the competent law enforcement agencies, equipped with the appropriate decryption conventions, can easily decipher the messages exchanged; an alternative consists in forcing the operators of communication platforms to provide the messages in clear text to the competent authorities upon request, which presupposes that these operators know said decryption conventions. Australia was the first democratic country to pass a law that forces cryptographic solution providers to provide mechanisms to circumvent encryption.[13] Intelligence agencies also operate at other levels: on the one hand, by individually tampering with equipment to weaken it (Gallagher, 2014) and, on the other hand, by proposing biased algorithms whenever they can (Green, 2013; Schneier, 2007), thus weakening all systems from players implementing these algorithms.

Yet, encryption weakening mechanisms are inherently inefficient and dangerous. Their limits were quickly revealed with implementation in the 1990s (Abelson et al., 2015). As soon as the existence of such flaws is known, users tend to turn away from the technologies concerned.[14] Moreover, at present, the free access to many cryptographic library source codes and to the necessary scientific knowledge allows any information technology specialist with a moderate scientific level to design their own encryption tools and to distribute them widely. The prohibition of encryption would then only penalize law-abiding citizens.

The fact that the underworld uses secure encryption does not prevent law enforcement from taking action. For example, through the compromission of the EncroChat encrypted communication system, law enforcement has been able to dismantle numerous criminal networks (Cox, 2020). The absence of preexisting backdoors is therefore not, as such, an absolute guarantee against interception. On the other hand, it does prevent such interceptions from scaling up, in accordance with Rule #2.3 of these specifications.

In some countries, where weakening encryption by coercing designers has not been considered a viable option, pressure has been put on users. In France, Article 434-15-2 of the Penal Code punishes the fact that "anyone who has knowledge of the secret decryption convention of a cryptology means likely to have been used to prepare, facilitate or commit a crime or an offence, refuses to hand over the said convention to the judicial authorities or to implement it, on the request of these authorities." While providers of technical means are clearly concerned, the article also explicitly targets defendants. The conformity of this article with the right not to incriminate oneself was therefore very quickly discussed.

The European Court of Human Rights (ECHR), in its decision in *Saunders v. United Kingdom*,[15] set out the contours of the right to remain silent in criminal matters, in particular by identifying the need to mobilize the suspect's "will" for the information to be revealed. However, the French Conseil Constitutionnel validated Article 434-15-2 of the Penal Code[16] on the grounds that encrypted data would exist outside the will of the individuals and that it would not therefore be a matter of forcing a confession but simply of accessing said data. However, the opposite of the right to remain silent is not the obligation to confess but to provide information that could incriminate the person or a third party. When an accused person withholds the location of a corpse, the corpse exists independently of the accused person's will, but this will is necessary for the location to be known. The same is true for numerical data: their meaning can only be revealed by the express will of the person concerned.

At a time when digital tools mediate more and more of our social interactions, such obligations to reveal encryption agreements are intended to force people to expose all of their words to the scrutiny of law enforcement authorities. Instead, fundamental rights should be guaranteed in digital spaces as part of a "digital habeas corpus."

Rule #3.1: Do not weaken by design the means of encryption.
Rule #3.2: Recognize the right to remain silent in the digital age.

Identities and Biometrics

The digital management[17] of identities is part of the process of rationalizing administrative practices of registration carried out in the West in the nineteenth century. The use of numbering, which began in the military and prison environments, was gradually extended to the entire population, with each person being assigned one or more registration numbers, sometimes from birth.[18] These practices have been complemented by the use of biometrics,[19] primarily in the judicial and law enforcement domains (Mattelart and Vitalis, 2014). Modern informatics has accelerated this trend by generalizing the use of digitized information to represent people.

Under the impetus of public authorities and pressure from the private sector, digital identity management systems are emerging in many countries. These are even being promoted as human development issues.[20] In the European Union, the eIDAS regulation[21] aims to establish a uniform regulatory and technological framework constituting a "common basis for secure electronic interactions between citizens, businesses and public authorities." In addition, a 2019 European regulation[22] now obliges member states to introduce identity cards with biometric data (fingerprints and photographs) stored in an embedded electronic component.

At the same time, the simplification work currently being carried out within administrations is promoting the doctrine of "tell us once", which aims to avoid users entering the same administrative data twice. However, in France, one of the effects of the 1978 Informatics and Liberties law was to abandon the use of a unique personal identification number, called the NIR, to uniquely identify individuals within the various administrations, as envisioned in the 1974 SAFARI project. The use of the NIR was thus restricted to the medico-social sphere, with other identifiers introduced in other sectors to deliberately make it difficult to cross-reference different administrative files. The current appetite of users for administrative operations that are as simple as the private services they use at the same time is encouraging the administration to abandon the compartmentalization that was intended to protect them and is leading to increased interconnections between the information systems of the various administrations.

The use of biometrics by states to strengthen the security of administrative identities raises additional questions. Biometrics has two main uses: authentication and identification. Authentication aims to determine whether a person is who they claim to be. This is done by comparing the person's biometric data with data that has been previously collected in a controlled manner from the person with that identity. If the data matches,[23] the person is indeed the one with this identity; if not, it is another person, with another identity, without knowing which one. Identification, on the other hand, aims to find the identity associated with a biometric trace that has been collected (for example, from a crime scene or taken from an unknown corpse). This trace will then be compared with all the biometric data contained in a reference database, associated with known identities, in the hope of finding a match. If such a match is found, the identity of the person who left the trace is revealed; if not, it means that the person is not present in the file. Performing an identification is therefore equivalent to performing an authentication with each of the entries in the database, hoping that one of them will succeed.

Identification systems therefore require a centralized database, whereas authentication systems can do without one. Indeed, in the case of authentication, the biometric identification elements can be stored on a medium "at the user's hand," such as a badge or an identity document embedding an electronic component. The person wishing to prove their identity simply presents the document they possess, whose biometric data is read and loaded locally on the comparison system, and then exposes the part of their body corresponding to the data on the document (face, finger, iris, etc.) so the comparison can take place. The fact that the user keeps their biometric elements at their hand offers many advantages in terms of computer security: it allows for the authentication of a person anywhere, by means of local systems able to read the data on the medium, without the need to open remote accesses to a centralized database that would be a prime target for computer attacks (Taylor, 2019).

However, some countries, when they issue biometric administrative documents, also keep the biometrics of applicants in a centralized administrative database, even though, as has been said, such storage is not necessary for the system to function. This is the case in France, which, for both the passports and identity cards it issues, stores these data in the “Titres Électroniques Sécurisés” (TES) database.[24] This storage has been strongly criticized for its risk of misuse.

Raymond Forni (2002), the very first vice president of the Commission Nationale de l’Informatique et des Libertés (CNIL), the French data protection authority, made a clear statement on the subject:

> In a democracy, I consider it necessary that there be room for fraud. If it had not been possible to manufacture fake identity cards during the war, tens of thousands of men and women would have been arrested, deported and probably killed. I have always been in favor of preserving a minimum of space, without which there can be no true democracy.

While it is not a question of allowing anyone to easily forge fake documents, it is a question of preserving, in the system of management of regalian administrative identities, the possibility for personnel to take the decision, in their soul and conscience, to create “genuine” fake documents. However, forging a “legendary” regalian identity out of nothing is no longer appropriate in this day and age because of the multiplicity of administrative files: for example, a person who has a regalian identity without having a history of care with a health insurance fund would easily be unmasked.

Plausible fake identities could therefore only be created by usurping the identity of existing or former persons. However, law enforcement authorities, having access to all the information in these files, would be much more likely to trap identity thieves because they can question them about multiple aspects of the lives of their doubles. Above all, keeping biometric information in databases such as TES would allow a tyrannical government to set up a centralized system of “de-duplication”; that is, to check whether, when a request for a document is made, the biometric elements provided do not correspond to a person already on file under a different identity.[25] The renunciation to the democratic postulate (Rule #1.1) therefore requires that the storage of such data be refused (Guo and Noori, 2021) and that citizens’ biometrics should always remain in their hands, save for police files. However, given the multiplicity of data that is now available on individuals, including openly on social networks, it is not certain that the ability to save lives by forging identities will be practicable on a large scale in a highly digitized future; the hope lies in the fact that disorders of sufficient magnitude are likely to undermine the integrity of the digital environment, reducing the risk of data cross-referencing by malevolent actors.

Rule #4.1: The establishment and management of identity documents shall be decentralized.

Rule #4.2: States shall not keep in their possession the biometrics of their citizens, save for police files.

Genetic Data

Genetic data is a special category of biometric data, and for this data, Rule #4.2 must be fully applied. Genetic data is extremely intrusive because it can reveal a person's past (direct and ancestral), present (likelihood of physical appearance, depending on the ability of certain genes to express themselves), and even probable future (via the susceptibility to the occurrence of a degenerative disease, for example). A person's genetic heritage is not their own: it is derived from that of their parents, will be partially transmitted to their eventual descendants, and is shared, in whole[26] or in part, by their (half) brothers and sisters and collaterals.[27] Heredity is encoded within DNA molecules that serve as the carrier of the genetic code. According to the current state of science, DNA is made up of "coding" and "noncoding" segments. While both types of segments are genetic data, they have different functions in cells. Coding DNA translates directly into proteins and regulatory mechanisms, which is apparently not the case for noncoding DNA.

Because genetic data is extremely intrusive, the legislators of the European Union and its member states have strictly regulated its processing. One of the safeguards implemented has been the distinction made between the coding and noncoding parts of DNA. Currently, in France, the law authorizes the storage and comparison of only certain well-defined noncoding segments, as well as the sex marker.[28] This choice is not insignificant: on the one hand, noncoding segments exhibit greater variability than coding segments and are therefore much more useful as discriminators between individuals; on the other hand, it was considered that coding segments, linked to the biology of the individual, were more intrusive and therefore did not need to be processed if the noncoding segments made it possible to discriminate individuals with sufficient reliability.

However, these safeguards are far from sufficient, as illustrated in France by the drift of the Fichier National Automatisé des Empreintes Génétiques (FNAEG). This file, created in 1998, was initially intended to store the genetic data of convicted sex offenders as well as unidentified genetic traces taken from crime scenes, to which new traces could be compared, as well as the "genetic fingerprints of persons against whom there is serious and corroborating evidence to justify their indictment for one of the offences" in question;[29] its planned number of samples was approximately two thousand. However, the scope of the FNAEG was soon extended, first to serious crimes and then to a range of offences, and it no longer concerned only convicted persons but also those for whom there were "plausible grounds for suspecting that they had committed an offense."[30] These combined changes led to a dramatic increase in the number of entries in the file, which, at the end of 2021, contained the fingerprints of 5,219,947 people (i.e., more than 7.5% of the French population) as well as 805,998 traces.[31] This manifest disproportion of the file led to a condemnation of France by the ECHR in 2017,[32] without the situation changing significantly.[33] It is worth noting that in 2015, of the 3,006,991 prints contained in the FNAEG at the time, 76 percent concerned defendants (Gueye and Pellegrini, 2017). This proportion is unlikely to decrease in the future as there is still no provision for the automatic erasure of persons who have not been convicted of an offence. The need to justify the request for deletion, following a discretionary maintenance decision by the public prosecutor, leads to a reversal of the burden of proof and the existence of an autonomous sanction for presence in this file outside of any proceedings against the individuals.

Although the Conseil Constitutionnel, in a 2010 decision, considered the principle of the FNAEG to be in conformity with the French

Constitution,[34] the extremely high proportion of persons who have not been convicted of an offence raises questions. It is evidence of a silent drift of the purpose of the FNAEG, which has become, because of the very broad collection practices and the low number of deletions, a stock of genetic data of "innocent persons" who have neither been convicted nor kept under investigation. Such a use is likely to contravene the limits set by the Conseil Constitutionnel in 2012 when it was asked about the proportionality of a proposed mega-ID file.[35] On that occasion, the Conseil Constitutionnel had censured the creation of a "file of honest people" that could be used for judicial purposes and would have contained a very large part of the population, which is the case here in view of the very large number of people kept in the file without having been convicted.

Another profound transformation in the use of the FNAEG occurred in 2016, when the legislature legalized the use of the FNAEG for direct lineage search.[36] The aim of kinship search is to identify, within the file, persons related to the bearer of the trace analyzed (Gershaw et al., 2011). The higher the degree of kinship and the weaker the statistical match, the larger the circle of potential suspects (up to several hundred or even thousands of individuals) who will then have to be eliminated by traditional investigation methods. This is why direct lineage search is the easiest to implement.

Research on relatives, used in a judicial context, profoundly changes the purpose of the FNAEG. It amounts to using this file as a stock of genetic material of persons recognized as not being directly concerned by the search carried out but whose collaterals may be implicated by ricochet. By applying a multiplier of 5.1 between the number of people present and those who can be implicated directly or through direct lineage (Gueye and Pellegrini, 2017), in 2021, this file included more than a third of the French population. By including siblings, the factor increases to 6.2, or more than 45 percent of the population. Indeed, the effectiveness of such a file for indirect kinship search increases with its size: the more a trace can be related to individuals from different families, the easier it will be to identify the person through a simple genealogical study. When applied to a large file, kinship search amounts to the implementation of an occult genetic file of the whole population.

The extension of this file, through the excessive retention of former defendants and the implementation of kinship search, makes it a mega genetic file of innocent people that does not respect the criteria set out by the Conseil Constitutionnel in its 2012 decision. The rights of the ascendants, descendants, and collaterals of the persons recorded in on file are also totally disregarded.

Although the law in France restricts the performance of DNA tests to three purposes: scientific, medical, and judicial, all of which are strictly regulated,[37] the commercial offer of DNA analysis by foreign private companies is not without its problems. In particular, American companies offer commercial services to people from all countries who submit samples to them to "determine their origins," but also to put them in touch with ascendants, descendants, and collaterals who have also undergone these tests, via the creation of huge genetic databases. However, these services and databases can easily be misused. People have used them to verify the paternity of their children and law enforcement authorities in several countries have submitted samples from crime scenes as if they were their own to obtain kinship information to guide their investigations. All of this contributes to the constitution, outside of any protective framework, of genetic databases of the world's population, which could easily be turned against the individuals and peoples that compose it.

> *Rule #5.1: States shall strictly limit the collection of DNA data to judicial purposes and for the most serious offences.*
>
> *Rule #5.2: The processing of genetic data retained in the judicial context may not be extended to kinship search.*
>
> *Rule #5.3: Leisure biometrics shall be prohibited.*

Facial Recognition

Facial recognition is a biometric technology that, once the morphological characteristics of a face have been processed to calculate a template, can be used either for authentication (by comparison with a single reference template stored on a medium, preferably at the user's hand) or for identification (by comparison with all the templates in a database of known individuals). This technology raises many questions in terms of necessity and proportionality insofar as, due to the existence of widely accessible data stocks (annotated photos on open social networks), the possibility of deploying these completely "contactless" systems without the knowledge of individuals and in high-traffic areas would make it almost impossible to avoid being subjected to them, especially when it is forbidden to conceal one's face in the public space.[38]

In the field of security, the techno-solutionist narrative presents augmented video for facial recognition in public spaces as a tool to effectively and cheaply identify wanted or undesirable people in certain areas (such as hooligans banned from stadiums). However, its practical effectiveness is questionable. Due to the intrinsic uncertainties in the production and comparison of templates, these methods inherently generate both false negatives and false positives. Even if their error rates are low, applying these methods to large populations leads to a large number of false positives (Telegraph Reporters, 2018). Based on the feedback from various experiments that have been carried out in some member states, the European Agency for Fundamental Rights (2019) published a summary document at the end of 2019 outlining the conditions for the lawful use of such systems in public spaces. In the United Kingdom, the experimentation carried out between 2017 and 2019 by the South Wales Police Force has been ruled illegal.[39]

Conceptually, if current facial recognition technologies are problematic because of their error rates, an errorless facial recognition technology would be even more problematic. It would lead to each person permanently displaying a unique and indelible barcode on their forehead that would distinguish them from others. This is a practice of totalitarian essence, violating Rule #1.2 and deeply damaging the fundamental freedoms guaranteed by a democratic society. Faced with the untenable technical promises of these devices and the proven risks for civil liberties, several European bodies have already taken a position in favour of a strict framework, such as the Advisory Committee of Convention 108 on the protection of personal data;[40] a moratorium, such as the European Parliament;[41] or even a pure and simple ban, such as the European Data

Protection Board, which brings together all of the European Union's data protection authorities.[42]

> ***Rule #6.1: Facial recognition in public spaces shall be prohibited.***

Conclusion

In contrast to the fascination with technology, which is fostered in governments and the public by promoters who are very often also sellers, these democratic specifications invite us to take the distance and time necessary for a reasoned assessment of the benefits (often immediate) and risks (often cumulative) of introducing technical innovations into the social field. The specifications listed here are far from exhaustive. This is a "version 1.0," intended to be enriched in the future, including by third-party contributions, to deal in depth with other uses of new technologies that bear major risks for democratic societies.

Several recent digital tools pose unprecedented civilizational risks, such as commercial social networks. The learning mechanisms deeply embedded in the structure of our brains, of an inductive and Pavlovian nature, lead to a coincidence being considered as a causality and a repetition as a certainty (Ripoll, 2020). By relying on these mechanisms and on our other cognitive biases, many platform operators participate for profit in the radicalization of opinions (Missika and Verdier, 2022). The result is deep and lasting fractures within human societies, ferments of conflicts, as shown by the attacks on the US Capitol on January 6, 2021, and on the Praça dos Três Poderes in Brazil on January 8, 2023 (see Blais, this volume; Oullet and Dufresne, this volume, for discussions regarding the use of individual data for microtargeting in the political sector). Whether personalized content filtering tools should only be regulated or banned is an open question, but one that needs to be addressed urgently. The same is true of the widespread use of inductive algorithmic processing, in my view wrongly called "artificial intelligence," such as text and image generating tools.[43] The capacity of such tools to produce content blatantly erroneous in substance yet with an elaborate form constitutes an unprecedented threat to the structure of human knowledge. The latter is threatened by a new form of censorship: not by scarcity, like the libraries burned in the past, but by abundance, by being submerged by wrong contents having the appearance of authenticity.

Faced with the risks of misuse of ever more powerful technologies, the intransigent application of the rules of the democratic specifications, regularly completed to take into account the scientific and technical advances, constitutes a duty towards the populations of our democratic regimes so that they can remain so as long as possible.

Notes

[1] This principle is called the "right of revolution" in the Anglo-American sphere.

[2] The confusion between these two notions (which have distinct names in French: *sécurité* versus *sûreté*) has misled even the CJEU (Tracol 2021, Section 6.4).

[3] A much better adjective than the misleading term "intelligent" sometimes used to characterize them.

[4] Ruling C-293/12 "Digital Rights Ireland" of April 8, 2014.

[5] Ruling C-293/12 "Digital Rights Ireland" of April 8, 2014, § 37.

[6] Such as the global analysis of telephone geolocation data having been conducted in France immediately after the November 13, 2015, attacks to discover possible accomplices of the attackers.

[7] Section 154 of Act no. 2019-1479 of December 28, 2019.

[8] One can notably mention, within the French administration under the Vichy regime, the role played by René Carmille, one of the first resistance fighters and martyrs of the digital age.

[9] This term refers to the ability to increase the range or power of a system at low cost.

[10] Such as the Privacy and Civil Liberties Oversight Board in the United States.

[11] See recommendation n. 31 of the Guide d'hygiène informatique v. 2.0, 42, published by the French ANSSI.

[12] Notably under Article 32(1)(a) of the GDPR, which explicitly encourages the use of data encryption.

[13] Telecommunications and Other Legislation Amendment (Assistance and Access) Act, no. 148, 2018.

[14] The "Clipper" chip implementing weakened encryption was a commercial failure because it found no export markets.

[15] ECHR, Saunders v. United Kingdom, req. no. 19187/91 [GC], 17/12/1996; see §74.

[16] Conseil Constitutionnel, decision no. 2018-696, March 30, 2018.

[17] Understood in its original sense of coding information in the form of numbers.

[18] This is the case in France with the NIR.

[19] By means of techniques such as "spoken portrait," photography, fingerprints, genetic markers, et cetera.

20 The United Nations' Sustainable Development Goal no. 16 ("Peace, justice and effective institutions") promotes target no. 16.9, which states: "By 2030, ensure legal identity for all, including through birth registration." In the name of this goal, the World Bank is financing digital civil registration programs in developing countries, which are huge markets for large companies in the sector.

21 Regulation (EU) no. 910/2014 of the European Parliament and of the Council of July 23, 2014.

22 Regulation (EU) no. 2019/1157 of the European Parliament and of the Council of June 20, 2019.

23 Since comparisons are made on templates containing reduced information and subject to margins of error, the comparison methods used are probabilistic by nature and therefore generate false positives and false negatives. They are therefore not fully reliable systems.

24 Decree no. 2016-1460 of October 28, 2016.

25 In France, while the architecture currently implemented within the TES file prevents a person's administrative identity from being retrieved from the biometric database, it would take very little time to modify it to add this feature.

26 In the case of homozygous twins.

27 Genetic data is an illuminating example of the inappropriateness of a "property" regime for personal data.

28 The comparison is authorized by article 706-56-1-1 of the Code of Criminal Procedure, which refers to a government order that provides the list of authorized noncoding segments.

29 Article 28 of law no. 98-468 of June 17, 1998, creating article 706-54 of the Code of Criminal Procedure.

30 Article 29 of law no. 2003-239 of March 18, 2003.

31 Document 5436/22 of the Council of the European Union of March 25, 2022, 16.

32 ECHR, Ayçaguer v. France, req. no 8806/12, June 22, 2017.

33 Decree no. 2021-1402 of October 29, 2021.

34 Conseil Constitutionnel, decision no. 2010-25 QPC, September 16, 2010.

35 Conseil Constitutionnel, decision no. 2012-652 DC, March 22, 2012.

36 Article 80 of Law no. 2016-731 of June 3, 2016, which creates Article 706-56-1-1 of the Code of Criminal Procedure.

37 Article 226-28-1 of the penal code punishes with a heavy fine persons who would carry out genetic analyses, with respect to themselves or third parties, outside this framework.

38 Law no. 2010-1192 of October 11, 2010.

39 EWCA, Bridges v. The Chief Constable of South Wales Police, no C1/2019/2670, [2020] EWCA Civ 1058, August 11, 2020.

40 Advisory Committee of the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (Convention 108), Guidelines on facial recognition, June 2021. See Debet (2021).

[41] European Parliament resolution of January 20, 2021, paragraph 56.

[42] Joint opinion 05/2021 of the EDPB and the EDPS, June 18, 2021.

[43] Which are in fact "stochastic parrots," see Bender (2021).

References

Abelson, A., Anderson, R., Bellovin, S.M., Benaloh, J., Blaze, M. Diffie, W., Gilmore, J., Green, M., Landau, S., Neumann, P.G., Rivest, R.L., Schiller, J.I., Schneier, B., Specter, M., and Weitzner, D.J. (2015). *Keys under doormats: Mandating insecurity by requiring government access to all data and communications*. Computer Science and Artificial Intelligence Laboratory Technical Report. Massachusetts Institute of Technology.

Bender, E.M., Gebru, T., McMillan-Major, A., and Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, 610–23.

Black, E. (2001). *IBM and the Holocaust*. Dialog Press.

Chavalarias, D. (2022). *Toxic Data*. Flammarion.

Cox, J. (2020, July 2). How police secretly took over a global phone network for organized crime. *Vice*. https://www.vice.com/en/article/3aza95/how-police-took-over-encrochat-hacked.

Debet, A. (2021). Reconnaissance faciale : un encadrement strict prôné par les lignes directrices du Comité consultatif de la Convention 108. *Communication commerce électronique*, *3*, 33–35.

Desprez, F. (2010). L'identité dans l'espace public: du contrôle à l'identification. *Archives de politique criminelle*, (1), 45–73.

Ellul, J. (1964). *The technological society*. Vintage Books.

European Union Agency for Fundamental Rights. (2019). *Facial recognition technology: Fundamental rights considerations in the context of law enforcement*. https://fra.europa.eu/en/publication/2019/facial-recognition-technology-fundamental-rights-considerations-context-law.

Gallagher, S. (2014, May 14). Photos of an NSA "upgrade" factory show Cisco router getting implant. *ArsTechnica*. https://arstechnica.com/tech-policy/2014/05/photos-of-an-nsa-upgrade-factory-show-cisco-router-getting-implant/.

Gershaw, C.J., Schweighardt, A.J., Rourke, L.C., and Wallace, M.M. (2011). Forensic utilization of familial searches in DNA databases. *Forensic Science International: Genetics*, *5*(1), 16–20.

Gormand, G., Malnory, C., Celle, P., Garcia, E. and Rozanac, E. (2021). Évaluation de la contribution de la vidéoprotection de voie publique à l'élucidation des enquêtes judiciaires. Centre de recherche de l'École des officiers de la Gendarmerie nationale, Grenoble Alpes Métropole. Final report of the study no 31300102.

Green, M. (2013, 18 September). The Many Flaws of Dual_EC_DRBG. *A few thoughts on cryptographic engineering* blog. https://www.schneier.com/blog/archives/2007/11/the_strange_sto.html.

Gueye, O., and Pellegrini, F. (2017). Vers une remise en cause de la légalité du FNAEG? *Convergences du Droit et du Numérique*, Forum Montesquieu, University of Bordeaux, 1–9.

Guo, E., and Noori, H. (2021, August 30). This is the real story of the Afghan biometric databases abandoned to the Taliban. *The MIT Technology Review*. https://www.technologyreview.com/2021/08/30/1033941/afghanistan-biometric-databases-us-military-40-data-points/.

Harari, Y.N. (2016, March 31). La stratégie de la mouche : pourquoi le terrorisme est-il efficace ?, *L'Obs*. https://bibliobs.nouvelobs.com/idees/20160331.OBS7480/la-strategie-de-la-mouche-pourquoi-le-terrorisme-est-il-efficace.html.

Mattelart, A., and Vitalis, A. (2014). *Le profilage des populations. Du livret ouvrier au cyber-contrôle*. La Découverte.

Missika, J.-L., and Verdier, H. (2022). *Le business de la haine. Internet, la démocratie et les réseaux sociaux*. Calmann-Lévy.

Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism*. PublicAffairs.

Mucchielli, L. (2018). *Vous êtes filmés ! Enquête sur le bluff de la vidéosurveillance*. Armand Colin.

Pellegrini, F., and Verdon, E. (2022). Les données à caractère personnel, carburant du capitalisme de surveillance. *L'Économie politique*, (2), 36–47.

Pellegrini, F., and Vitalis, A. (2017). La création du fichier biométrique TES : la convergence de logiques au service du contrôle. *Sociologie*, 8(4), 447–52.

Ripoll, T. (2020). *Pourquoi croit-on ? Psychologie des croyances*. Éditions Sciences Humaines.

Rouvroy, A. (2014). Des données sans personne : le fétichisme de la donnée à caractère personnel à l'épreuve de l'idéologie des Big Data. In *Le numérique et les droits et libertés fondamentaux*. Étude annuelle du Conseil d'État.

Schneier, B. (2007, 15 November). The Strange Story of Dual_EC_DRBG. *Schneier on Security* blog. https://www.schneier.com/blog/archives/2007/11/the_strange_sto.html.

Singh, S. (1999). *The Code Book*. Doubleday.

Sureau, F. (2022, June 3). Discours de réception de M. François Sureau. Académie française. https://www.academie-francaise.fr/discours-de-reception-de-m-francois-sureau.

Taylor, J. (2019, August 14). Major breach found in biometrics system used by banks, UK police and defence firms. *The Guardian*. https://www.theguardian.com/technology/2019/aug/14/major-breach-found-in-biometrics-system-used-by-banks-uk-police-and-defence-firms.

Telegraph Reporters. (2018, May 5). Police defend facial recognition technology that wrongly identified 2,000 people as potential criminals. *The Telegraph*. https://www.telegraph.co.uk/news/2018/05/05/police-defend-facial-recognition-technology-wrongly-identified/.

Tracol, X. (2021). The two judgments of the European Court of Justice in the four cases of Privacy International, La Quadrature du Net and Others, French Data Network and Others and Ordre des Barreaux francophones et germanophone and Others: The Grand Chamber is trying hard to square the circle of data retention. *Computer Law and Security Review*, *41*, 105540.

Tracol, X. (2023). The joined cases of Dwyer, SpaceNet and VD and SR before the European Court of Justice: The judgments of the Grand Chamber about data retention continue falling on deaf ears in Member States. *Computer Law and Security Review*, *48*, 105773.

Zuboff, S. (2019). *The age of surveillance capitalism*: *The fight for a human future at the new frontier of power.* Public Affairs.

This page intentionally left blank

PART 2

Politics: Use of Data, Profiling, and Personalizing in Electioneering

This page intentionally left blank

The Closing of Ranks

4

The Collusion of Federal Political Parties and the Resistance to Privacy Regulation

Colin J. Bennett

Introduction

Canadian political parties are not generally covered by Canadian privacy legislation – either at the federal or provincial levels.[1] A ten-year campaign to bring them under the umbrella of privacy law, and the oversight of Canada's Information and Privacy Commissioners, has been met with stiff and generally unified resistance. After consistent pressure, political parties are still the one category of organization in Canada over which individuals have few, if any, legal privacy rights.

Canadian parties obviously operate in a highly competitive political environment and have had to adapt to widespread changes in communications technologies and practices. However, they are also entrenched and prone to collectively defend their interests against regulators. Indeed, there is some literature that suggests they can operate as a form of "cartel" and have colluded in the past to exclude new parties from obtaining official party status, to shape the provision of state financial subsidies to benefit their own interests, and to regulate ballot access (Gauja, 2014; MacIvor, 1996; Young, 1998). A significant resource for the modern digital campaign is personal data. Privacy regulation curtails the freedom of organizations to collect, process, and disclose personal data without consent. Thus, can the resistance to privacy regulation by federal political parties be explained by this same "political party cartel" theory? And if so, how does the cartel operate, and how is the collusion observed?

In this chapter, I review the multipronged strategy to bring political parties within the ambit of Canadian federal privacy law over the last decade. I also document the parliamentary pressure, litigation, targeted complaints to regulators, and public and media advocacy, record the various statements by party officials and spokespeople, and critique the various stated justifications for treating parties as exceptional organizations. I contend that none of the parties' arguments stands up to serious scrutiny. There is no plausible reason why the federal political parties (FPPs) should not have to abide by the same standard privacy regulations as businesses, public bodies, and government agencies.

There is a deeper purpose to this analysis beyond concern about the privacy of Canadian voters. The processing of personal data on the electorate permits practices of microtargeting, with some profound implications for democratic discourse and engagement (Bennett and Gordon, 2021) and for the deliberative quality of Canadian democracy (see Blais, this volume). The regulation of the conditions under which personal data might be collected, processed, and distributed by, and for, political parties also permits us to get a handle on some of the more nefarious practices in Canadian election campaigning, particularly online. I contend that a deeper understanding of the resistance to privacy protection in the political sphere is also key to rendering contemporary data-driven elections more transparent and thereby combating some of the worst effects of contemporary digital campaigning.

Political Parties, Cartels, and Collusion

Evidence of collusion between different political parties of different persuasions and ideologies has been found in many societies. It is behaviour most notably associated with the theory that dominant parties can often display "cartel-like" behaviour when they collectively believe their interests are threatened.

The notion of political parties as cartels emerged out of a deep-seated concern with the decline of political parties in Western democracies and the "dealignment" of Western electorates. Richard Katz and Peter Mair (1995) argued against the prevalent model of the "mass-party" based on predefined social groupings. The party, in this model, represents the interests of these groups in its program and is dependent on these groups for membership and financing. Electoral success, according to this model, turns on the parties' abilities to mobilize their support. It is this model of the political party, and of their linkages with civil society, that was supposed to have atrophied in the 1960s and beyond as parties became increasingly delinked from traditional bases of support (Dalton and Wattenberg, 2002). Yet, Katz and Mair (1995) contend that this model was never really predominant and was based on an idealistic conception of the dependence of political parties on civil society. Just as important, they contend, is the relationship between political parties and the agencies of the state. They coined the term "cartel party" to describe a more realistic conception of modern parties and their development and to recognize that parties have organizational interests that may be distinct from those of the social groups that support them.

The concept of the cartel party, and the theorizing it inspired, motivated a great deal of debate. It is posited as a quite general theory, applicable to many modern democracies. It was also, of course, developed before contemporary practices of digital campaigning became commonplace. Nevertheless, its central claims may still have considerable relevance. A critical component of the theory is the emphasis on interparty cooperation and collusion. As the same authors remarked in a later defence of their thesis, the depoliticization of the electorate and the dealignment of broad social groups from the main parties makes it easier for parties to cooperate and collude across a wide range of policy issues (Katz and Mair, 2009).

There are a number of empirical studies that have documented the weakening of parties' linkages with social groups and their intensification of relations with the state. In a crossnational analysis of European political parties, Ingrid van Biezen and Petr Kopecký (2014) identified three dimensions of state-party linkages: the dependence of the parties on the state, the management of parties by the state, and the capture of the state by parties. Much of their analysis focused on the increasing reliance by parties on the state for financing. That dependence has been accompanied by increasing regulation of parties' activities and, in several countries, gives the major parties a quasi-official status concomitant with their institutional relevance. That status contributes to cartel-like behaviour, with explicit and latent attempts to exclude other parties from competition in the electoral arena.

In an analysis of political parties in common law democracies, Anika Gauja (2014) documented a wide range of party regulation to enforce the rules of "legitimate competition" and particularly rules for ballot access and registration requirements. States seek to control the number of political parties that participate in elections to reduce the complexity and fragmentation of the democratic process. These rules resonate with the party cartel thesis, even though their implementation in different countries is variable and does not always and unequivocally support cartelization. The courts also, Gauja argued, have important roles in striking the correct balance between constitutional rights to political speech and organization and the management of electoral competition.

However, the party cartel thesis has had less traction in the Canadian context. Heather MacIvor (1996) subjected the Katz and Mair party cartel thesis to empirical inquiry in the early 1990s. She found that the three main parties did indeed collude to exclude new parties from access to the ballot and used state subsidies to reinforce their own interests. But there was also a backlash against this behaviour, reflected in the rise of the Western-based Reform Party in the early 1990s. Thus, MacIvor concluded that the entire party cartel model is not really reflected in the historical development of party competition in Canada. Party cartels are inherently unstable, she argued. The model embodies significant contradictions that can challenge the self-protective behaviour of the major parties.

An article published around the same time by Lisa Young (1998) argued that there was little evidence that Canadian parties had transformed themselves into cartel parties, whose fundamental desire for self-preservation leads to a dependency on the state. Canadian parties had not become reliant on state resources, and their ties to groups and supporters in civil society remained strong. She goes further, arguing that the cartel thesis was far more applicable to European party systems than it is to Anglo-American democracies. Although there was plenty of evidence of collusion among political parties to defend their collective interests, this did not lead to the kind of overall cartelization that Katz and Mair had in mind.

Much of this literature dates from an earlier period of party competition and campaign practices. The party cartel thesis certainly drew attention to the increasingly managerial and professional, self-referential and technocratic character, of modern parties (Katz and Mair, 2009). With the overall decline in mass membership, there is no surprise that political parties would become more dependent on state subsidies and would sometimes collude to protect those assets.

So does this theory have any relevance in contemporary circumstances? In an era of digital campaigning, an increasingly significant asset for political parties is "big data" (see Trudel, this volume) from which inferences about political behaviour might be derived. Privacy protection law threatens the control over that asset. Do parties collude, therefore, to collectively protect those resources from outside regulation? Do they collectively defend their rights to capture, process, and disseminate personal data? In Canada, at any rate, there is considerable evidence that they do.

Information, Privacy, and Canadian Political Parties: The Regulatory Gap

In most democratic countries, the opportunities for political parties to capture and use personally identifiable data to identify and target voters are severely constrained by comprehensive privacy or data protection laws. Those laws define information on "political opinions" as highly sensitive and typically require the express consent of the individual before it can be processed (Bennett and Oduro Marfo, 2019). This is not the case in Canada. Political parties in Canada, like in the US and Australia and unlike in Europe, are generally viewed as outside the scope of federal or provincial information privacy laws. Therefore, the extent to which candidates and parties abide by the commonly agreed principles of information privacy protection has largely been a matter of choice rather than compulsion. Canadian political parties generally self-regulate and have produced various codes of practice and privacy policies designed to offer some assurances to Canadian voters and to convince outsiders that they do not need to be further regulated (Bennett, 2018).

Thus, for the most part, individuals have no legal rights to learn what information is contained in party databases, to access and correct those data, to remove themselves from the systems, or to restrict the collection, use, and disclosure of their personal data. Parties typically have no legal obligations to keep that information secure, to only retain it for as long as necessary, and to control who has access to it (Bennett and McDonald, 2020).

Canadian privacy protection law is a typical patchwork of federal and provincial law containing different provisions for public and private sectors. Privacy protection is such an example of a legal regime that has emerged pragmatically at both federal and provincial levels. Federal public sector agencies are governed by the Privacy Act, and there are separate laws that regulate the processing of personal data by provincial public bodies, most of which also offer access to information rights, such as those in Ontario and British Columbia (BC). Federally regulated private sector industries (banking, transportation, and communications), as well as companies that transmit personal data across international and interprovincial borders for commercial purposes, are governed by the Personal Information Protection and Electronic Documents Act (PIPEDA). Similarly, there are a handful of substantially similar provincial laws (in Quebec, BC, and Alberta) that govern the private sector in those provinces, including the BC Personal Information Protection Act (PIPA), discussed below. The Office of the Privacy Commissioner of Canada oversees compliance with the Privacy Act and PIPEDA, and similar offices exist in each province to oversee compliance with the equivalent provincial legislation.

So the vast majority of personal data captured by public and private organizations is covered by privacy protection law at some level. Yet the patchwork is complicated, not entirely comprehensive, and sometimes mired in interjurisdictional confusion. Canada's privacy protection regime depends on an uneasy distinction between the commercial and governmental processing of personal data. Political parties are a hybrid, and they tend to fall between the cracks of this complex patchwork of privacy laws (Bennett, 2018; Bennett and Bayley, 2012).

There are two provinces where provincial political parties are regulated under provincial privacy laws. In BC, PIPA covers "organizations" (other than public bodies) regardless of whether they are engaged in commercial activity. The Office of the Information and Privacy Commissioner of BC (OIPCBC), therefore, has jurisdiction over political parties and has already conducted three investigations. The first, involving the BC New Democratic Party (BC NDP), and the second, involving the BC Liberals, served to establish that the OIPCBC did indeed have jurisdiction in this area at least at the provincial level (OIPCBC, 2011, 2013). Those precedents led to a third broader analysis of compliance with PIPA by all major political parties in BC (OIPCBC, 2019a). The report concluded that BC political parties needed to be more transparent about how they collect data on voters; too much was being gathered without the individual's consent. The major value of this report is its portrayal of the range of data captured by Canadian political parties and of the various practices that have gradually crept into the campaigning culture.

A similar picture is provided by a report on Quebec elections by Élections Québec (2019), the independent body charged with regulating elections and campaigning procedures in Quebec. As elsewhere, Quebec parties leverage the data from the electoral lists to build databases to determine their various persuasion strategies and to personalize the messages delivered to voters. Élections Québec therefore recommended that both provincial and municipal political parties should be brought within the general legal framework for the protection of personal information. And in the recently enacted amendments to Quebec's personal data protection regime (Bill 64), provincial political parties are indeed covered. Section 127.22 states that the private sector legislation does apply to political parties, candidates, and members of the assembly, and that they "may not collect or use personal information without the consent of the person concerned." However, as explained below, these new rights are circumscribed as a result of strong pressure from Quebec political parties. There is general agreement that they do not match the requirements imposed on businesses or government agencies (Quebec National Assembly, 2021).

In summary, the lack of a general regulatory framework for Canadian political parties has allowed Canadian elections to be influenced by trends in voter analytics from south of the border (Bennett, 2015). These practices include the widespread use of voter relationship management (VRM) platforms, such as Liberalist or the Conservative's

Constituency Information Management System; the increasing use of mobile applications for household canvassing; close associations with voter analytics companies and consultants; and the increasing use of social media (mainly Facebook) to analyze issue trends and reach out to precise segments of the electorate through "microtargeting" (Bennett and Gordon, 2021; Delacourt, 2016; Dubois and Taylor, 2019).

Privacy regulation is not just about privacy rights, therefore. And it is not just about the technical and managerial questions about personal data protection. The various risks to privacy, civil liberties, and to democratic practices can be extensive when US-style campaigning techniques are imported into a country with a very different party system, electoral process, financing rules, and political culture. Comprehensive privacy protection rules also challenge the liberal use of consumer lifestyle data for political campaigning (see Ouellet and Dufresne, this volume). Thus, the appropriate protection of personal data is a question that goes to the heart of contemporary debates about democracy. National and transnational privacy protection laws are now at the centre of an international conversation about the erosion of democracy through the manipulation of big data on the electorate and the targeting of messages (true and false) to increasingly narrow categories of voters (Bennett and Oduro Marfo, 2019; Council of Europe, 2021).

Privacy Advocacy and the Campaign for Regulation

It took several years, and several scandals, for the weaknesses of privacy regulation in the political arena in Canada to be fully appreciated. A scandal concerning the highly controversial practice of robo-calling and voter suppression during the 2011 election was one of the first to draw attention to the regulatory gap. The scandal contributed to the heightened publicity about privacy and security protections for the parties' VRMs and led to a series of (unheeded) recommendations by the federal regulator of elections, Elections Canada, to apply basic privacy principles to Canadian political parties (Elections Canada, 2013).

At around the same time, the former privacy commissioner, Jennifer Stoddart, was receiving complaints about parties' practices and could not respond as she believed she did not have jurisdiction. That realization led to a report on privacy protection and Canadian FPPs (Bennett and Bayley, 2012) and to further public debate about the extent and nature of the problem. Several academic articles also analyzed the different sets of privacy questions raised when personal data is processed for the purposes of political communication and democratic engagement (Howard and Kreiss, 2010; Bennett, 2013; 2015; Rubinstein, 2014; Judge and Pal, 2021). The analysis of new practices of digital campaigning in the United States served as important warnings for other countries about the role that data analytics was playing in influencing voters in American elections (Issenberg, 2013).

The full extent of voter surveillance was not fully appreciated until the scandal involving Cambridge Analytica and its illegal harvesting of data from Facebook catapulted to global attention in 2017 (Cadwalladr, 2017). This story, of course, had everything to capture public and political attention for many months: colourful characters, massive implications for the outcomes of elections in the UK and the US, powerful and bold investigative journalism, shady Russian involvements, high-profile parliamentary inquiries, and the potential intervention of different regulatory authorities in several countries. The subsequent investigations in the US (Federal Trade Commission, 2019) and the UK (Information Commissioner, 2018) led to substantial fines for Facebook and the eventual bankruptcy of Cambridge Analytica.

The scandal highlighted a reality about contemporary elections that some had been warning about for many years. Election outcomes can be substantially driven by the analysis of personal data. Political elites had bought into the assumption that elections can be won if they just had more precise data on the electorate, through which they can profile voters and microtarget precise messages to key groups, either to encourage or suppress their participation.

There was also a substantial Canadian connection to the scandal. Chris Wylie (2019), the principal whistleblower, came from BC. The software company (Aggregate IQ) responsible for the aggregation and analysis of data on behalf of the Brexit campaign, which was subsequently fined for the illegal processing of personal data, is based in BC's capital, Victoria. And a prominent regulator, Elizabeth Denham, UK information commissioner, was already informed of the risks of digital campaigning in her former role as BC information and privacy commissioner. The

BC commissioner from 2018 to 2024, Michael McEvoy, was brought over to the UK to assist with the various investigations in privacy and political campaigning (ICO, 2018, 2019). These connections ensured that the various scandals were prominently reported in the Canadian media.

The rationale for the exemption of political parties from privacy protection laws was then more widely questioned and a broader campaign for legal reform emerged. Parliamentary oversight centred on the work of the House of Commons Select Committee on Access to Information, Privacy and Ethics (ETHI). After a series of hearings into the vulnerabilities of Canada's democratic system, in the wake of the breach of personal data involving Cambridge Analytica and Facebook, the committee recommended "that the Government of Canada take measures to ensure that privacy legislation applies to political activities in Canada, either by amending existing legislation or enacting new legislation" (ETHI, 2018). These hearings also offered an opportunity for the major parties themselves to be heard on the question and for outsiders to hear the various justifications for *not* subjecting the parties to regulatory oversight, discussed below.

At their 2018 annual meeting, the federal, provincial, and territorial information and privacy commissioners called on their respective governments to pass legislation (OPC, 2018):

- Requiring political parties to comply with globally recognized privacy principles;
- Empowering an independent body to verify and enforce privacy compliance by political parties through, among other means, investigation of individual complaints; and,
- Ensuring that Canadians have a right to access their personal information in the custody or control of political parties.

In April 2019, the privacy commissioner and the chief electoral officer issued joint guidance (OPC, 2019) on the protection of personal information in response to changes in the Elections Modernization Act of December 2018 (Bill C-76). This law obliges parties to have a publicly available, easily understandable policy describing the collection, protection, and sale of personal information, procedures for staff training, and the identity of a designated person to whom privacy concerns can be addressed. The submission of this policy is a condition of registration with Elections Canada. These provisions have been met with almost universal criticism for their incompleteness, vagueness, and lack of any real enforcement mechanism (Scassa, 2018). In consultation with Elections Canada, the former privacy commissioner recommended amendments and, in particular, a specification that the privacy policies must be consistent with the principles within the Model Code for the Protection of Personal Information, found in Schedule 1 of PIPEDA, and that their office should be responsible for oversight (Therrien, 2018).

Civil society organizations also exerted pressure. Most notably, Vancouver-based OpenMedia launched a petition and letter-writing campaign. They also developed a new tool (My Political Data) empowering people to make access to personal information requests from political parties. They reported over 1,700 requests for personal data using this tool. They also conducted a systematic analysis of the FPP's privacy policies in the aftermath of the guidance by the Office of the Privacy Commissioner of Canada (OPC) and Elections Canada. None, according to OpenMedia, came close to meeting the stated standards.

Public opinion surveys also registered both a general surprise that parties were not covered by our privacy laws and strong support for bringing them within the regulatory framework. A May 2018 poll from Innovative Research Group found that 37 percent of Canadians (from all political standpoints) believed that parties should be subject to stronger protections for personal information than private companies, and a further 49 percent believed they should be subject to the same laws. Another poll from the Campaign Research Group found that only 9 percent of the sample were aware that the parties were exempt from Canadian privacy protection laws. Sixty five percent strongly agreed that they should be (Curry, 2019). Strong empirical support for the scope of the problem was provided in a poll prepared for Elections Canada of candidates in the wake of the 2019 federal election. Candidates were asked about the steps they took to protect the voters list. Only 36 percent ensured the lists were kept secure and locked. Only 24 percent ensured the destruction of the lists at the end of the election campaign (Ekos Research Associates, 2020).

Thus, there is widespread political support for bringing political parties into the ambit of Canada's privacy laws. The key regulators made the appropriate recommendations. Media attention increased. In light of the overwhelming political reality, however, that a campaign to restrict parties' uses of personal data would also curtail the use of a key resource that brought them to power, very little has happened. Facing such challenges, well-argued advocacy and recommendation had been insufficient and stronger approaches were seen as necessary.

Complaints and Litigation

The Centre for Digital Rights (CDR) was established in 2018 by Jim Balsillie, former chair and co-CEO of Research In Motion (BlackBerry) and founder of the Balsillie School of International Affairs at the Universities of Waterloo and Wilfrid Laurier and of the Centre for International Governance and Innovation. Deeply concerned about digital rights and the impact of surveillance capitalism, he launched CDR as a not-for-profit organization to "promote public awareness of digital rights issues related to the digital economy" and advance "best practices, laws and regulations that protect the civic values and the rights of individuals in the 21st century economy driven by the mass collection, use, control and disclosure of data" (CDR, n.d.).

In Spring 2018, and in light of the fallout from the Cambridge Analytica scandal, CDR launched several legal complaints against the personal data protection policies and practices of Canada's three main FPPs. Those complaints were submitted to the Federal Commissioner of Competition under the Competition Act's prohibitions against deceptive marketing; the Chief Enforcement and Compliance Officer of the Canadian Radio-Television and Telecommunications Commission under Canada's Anti-Spam legislation; the Federal Privacy Commissioner under PIPEDA; the Commissioner of Elections Canada under the Canada Elections Act; and the Information and Privacy Commissioner of BC under PIPA. Three, in particular, require more detailed analysis. They also inspired some serious and collective resistance from the FPPs.

To the OPC, the CDR argued that FPPs have been subject to PIPEDA all along and that it was well within the OPC's jurisdiction to investigate their practices. PIPEDA applies to every organization with respect to the personal information "the organization collects, uses or discloses in the course of commercial activities" (s.4(1)(a)). The complainant was Mr. Robert (Gary) Dickson, former MLA in Alberta and Saskatchewan Information and Privacy Commissioner. The FPPs are engaged in commercial activities, and always had been, alleged the complaint. They sell party-branded memorabilia. They solicit donations. They engage the services of consultants and data analytics companies to promote their messaging. They purchase the services of social media companies. In short, they promote and protect their brands just as companies do. PIPEDA has a broad and quasi-constitutional status, and there was no intention to exclude FPPs from its application (OPC, 2021). In short, organizations do not need to be for-profit commercial entities to engage in commercial activity.

Both the NDP and the Liberal Party pushed back (the Conservatives did not engage). Both contended that they were not engaged in commercial activities within the meaning of PIPEDA. Further, they are subject to regulation under the Canada Elections Act (CEA), which includes requirements for the publication of privacy codes of practice and protections on the use and disclosure of the voters list. Their purpose is to promote participation in public affairs, a purpose that is inherently noncommercial. They also asserted that they did not "sell, barter, or lease its donor/supporter lists."

Federal Privacy Commissioner Daniel Therrien took many months to conclude his review of the complaint. In the commissioner's published response, he noted that Parliament did not expressly exclude FPPs from the ambit of the legislation and so PIPEDA could apply to the extent that the parties engage in "commercial activities" – such as the selling or purchasing of lists for exchange. However, he was not convinced that the political parties were engaged in commercial activities, under the ordinary meaning of that term. Further, the commissioner concluded that just because parties might be engaged in some commercial activity does not serve as an authorization to subject their entire operations to oversight. He also noted that Parliament had recently declined to make the FPPs subject to the act, preferring instead to amend the CEA with the requirement for self-administered privacy policies. The commissioner has advocated that PIPEDA apply to the FPPs, but he had to interpret and apply the law as written and not as he would like it to be written (OPC, 2021). The complainants did not seek a judicial review of the commissioner's decision.

With PIPEDA out of the picture, therefore, what of the application of the CEA? The focus of the complaint under the CEA was the use of the voters list. Section 109 of the CEA requires the Chief Electoral Officer to prepare final lists of electors to each registered party in an electoral district. Section 110(1) authorizes the registered FPPs to use the list for the purposes of "communicating with electors, including using them for soliciting contributions and recruiting party members." Section 111(f) prohibits the parties or candidates from knowingly using the personal information included in the list of electors for any other purpose. These provisions provided leverage for the CDR to complain to the Commissioner of Canada Elections (CCE) that indeed the list of electors was being used for other purposes, knowingly and in contravention of the CEA.

In my affidavit to the court in this case, I identified three reasons to support the complaint (Bennett, 2020). First, the identifying information from the list of the electors is the indispensable backbone of the increasingly sophisticated databases operated by the main FPPs. There was no contemplation at the time of the passage of the modern CEA in 2000 that the personal information from the list of electors would be amalgamated with other sources of data to construct these databases. It is reasonable to question whether "communicating with electors" also authorizes the building of databases on the entire electorate for profiling purposes, especially as such systems are the exception, rather than the rule, in the advanced democratic world.

Second, these systems, like any complex database, are subject to what is generally called "mission," "scope," or "function" creep. The expansion of a system or technology beyond its original purposes is a widely observed phenomenon in the study of science and technology and in the analysis of surveillance (Koops, 2021). The most telling example of this phenomenon is a media report that the Liberalist database was used to vet potential judges (Leblanc and Cardoso, 2019). More troubling was the testimony of a whistleblower, who had worked as a data analyst for each of the main parties, that these aggregated data have been used to damage political rivals (Newman, 2019).

Third, what does it mean exactly to "communicate *with* electors"? Times have changed radically since the CEA was passed. The extensive and global publicity about the use of voter analytics in campaigning has revealed more opaque and manipulative campaigning techniques that may fall outside the meaning of "communicating with electors" established in the act. If there are concerns that personal data included in the FPPs' databases, which in turn are based on the data gleaned from the list of electors, is being used not to "communicate *with* electors" but for more manipulative purposes, then a CCE investigation is warranted.

A key to an understanding of the difference between communication and manipulation is the technologies' and the data's predictive capacities, used to "nudge" behaviour in one direction or another. Thus, the party or candidate is not engaged in communication *with* the voter but in a process that taps into the unconscious to "create" voters – or nonvoters. The voters in this model are not intended to process and consider information and make an autonomous choice. In fact, the goal is generally the opposite, to appeal to implicit biases and vulnerabilities. There is little attempt to communicate *with* the voter and respect that voter's autonomy to make a democratic choice (Burkell and Regan, 2019).

In its response in April 2019, the CCE noted that its jurisdiction was limited to the relatively narrow category of information contained in the list of electors – name, address, and the unique voter number. The commissioner had no jurisdiction over the broader uses and potential misuses of personal data by political parties. The CDR sought a judicial review of the decision, contending that the commissioner had erred in his interpretation of the act. That review was unsuccessful. The decision did, however, clarify that there was no independent oversight over the personal data collected, processed, and disseminated by the FPPs beyond that contained in the list of electors. This was confirmed in the commissioner's 2021 annual report, in which he confirmed that "under the current legislative framework, very little can be done when electors come to us and express concerns or submit complaints about the use of their personal information by political parties" (Canada Commissioner of Elections, 2021). Thus, the parties could not credibly contend that their data operations needed no further regulation because they were already subject to a strict regulatory regime under the CEA. That issue came squarely to the fore with respect to the next set of complaints in BC.

PIPA was passed in 2004 as a private sector data protection act, substantially similar to PIPEDA. Its scope is not confined to commercial purposes. Rather, it covers all "organizations" that collect, use, and disclose personal information, unless explicitly exempted in the legislation. Political parties are not exempted, and provincial political parties, at least, are covered by PIPA and have been investigated, as noted above. Are FPPs also regulated, however, to the extent that they collect personal information in BC on BC residents?

This issue was initially joined as a result of a complaint in 2018 to the Courtney-Alberni Riding Association of the New Democratic Party of Canada. The complainants had received a personal email invitation to a meet-and-greet with NDP leader Jagmeet Singh and questioned how the NDP obtained their information. Citing PIPA, they sought to obtain information from the NDP about the extent of information collected, to whom it had been shared, and who might have had access to it. The individuals received no response and then complained to the OIPCBC. In response, the NDP took the position that PIPA does not apply to the federal NDP and that the BC commissioner had no jurisdiction (OIPCBC, 2019b).

Three constitutional questions were raised: PIPA is ultra vires by virtue of Sec. 41 of the Constitution Act; PIPA is inapplicable to political parties because the CEA, PIPEDA, and other federal laws are paramount; and PIPA amounts to unjustified limits on rights to vote and freedom of expression under the Charter. Commissioner McEvoy ruled against the NDP on the first two questions and did not reach a judgment on the Charter question. According to this ruling, therefore, a local "electoral district association" registered under the CEA is an "organization" and covered by PIPA (Scassa, 2019). The matter could then proceed to the merits of the complainants' allegations. The NDP did not appeal this order.

In the meantime, the CDR had been pursuing the issue further on behalf of three BC citizens who had asked the Liberals, Conservatives, NDP, and Greens for the personal information those parties held on them. The parties provided some information, but the complainants believed that the responses were neither complete nor adequate.

Supported by the CDR, they then appealed to the BC commissioner, who opened an investigation. The NDP, Liberals, and Conservatives then responded that the commissioner did not have jurisdiction because PIPA did not apply to them. The commissioner then appointed David Loukidelis QC, the former information and privacy commissioner of BC, to conduct a written inquiry as a delegate adjudicator to determine whether PIPA can constitutionally apply to FPPs.

What is most interesting about this dispute is how strenuously the three parties continue to fight the case. Each hired some very high-powered law firms to push back on the commissioner's assertion of jurisdiction. They each gave notice under the Constitutional Questions Act to the attorney generals of Canada and of BC. The Conservatives also wanted other federal parties given notice and for the privacy commissioner of Canada, and the chief electoral officer to be invited to intervene (OIPCBC, 2022, 9).

Each of the parties raised somewhat different objections. The NDP asserted that the province's entire law was essentially an unconstitutional expression of provincial power. The Liberals argued that they were not "organizations" under the meaning of the law. The parties also advanced a couple of broader constitutional arguments – provincial law actually frustrates "paramount" federal laws and particularly the CEA and impairs the federal government's exclusive power over federal elections and the federal government's interests in enforcing a uniform set of rules. The NDP even contended that the provincial privacy law could not apply because they were subject to the federal privacy law (PIPEDA). This argument was disingenuous. They are not, and they know they are not, because the federal privacy commissioner had already said so.

In response, Mr. Loukidelis rejected the parties' arguments, thereby providing the first unequivocal victory for the CDR and the complainants. His order of March 1, 2022, determined that FPPs are "organizations" within the meaning of PIPA; the application of PIPA is not ousted by the application of the federal privacy legislation; the essential "pith and substance" of PIPA, the regulation of the collection, use, and disclosure of personal information, is a valid exercise of provincial legislative authority respecting property and civil rights; and PIPA is not inapplicable to the FPPs by virtue of the constitutional doctrines of either paramountcy or interjurisdictional immunity. The issue of whether PIPA unconstitutionally infringes on the right to vote or freedom of expression as guaranteed in the Charter was not addressed (OIPCBC, 2022).

The argument that privacy protection law could have a "chilling effect" on political communication and participation has been advanced in the US from time to time, and reform faces massive hurdles in light of the presumptive significance of political speech under the First Amendment (Rubinstein, 2014). These arguments are somewhat new in Canada, however, and have been advanced in particular by the NDP in the various proceedings in BC. The essence of the argument is that political speech is restrained and made less effective if the parties know less about the individuals and groups receiving the messaging. Effective political communication in the modern era therefore requires detailed knowledge of the beliefs, interests, and likely behaviours of the audience. According to this hypothesis, to the extent that parties are restricted in gaining that knowledge by privacy regulation, their essential roles in engaging the electorate are restricted. Mr. Loukidelis did not reach a judgment on these wider considerations under Section 2 of the Charter of Rights and Freedoms in the absence of further evidence and a fuller briefing. The arguments are also, of course, open to the obvious counter that strong privacy protection can enhance the trust in the democratic process and therefore foster democratic engagement and participation (Cohen, 2021).

This was a lengthy decision, fully and extensively briefed by both sides. In April 2022, the Liberals, Conservatives, and NDP filed a judicial review of Mr. Loukidelis' order. The NDP and Conservatives both relied on the Liberals to lead the appeal and joined in the principal arguments.

After several delays, the case was heard in the BC Superior Court in April 2024, and Supreme Court Justice Gordon Weatherill quickly issued his decision on May 14, 2024 (BC Supreme Court, 2024). He affirmed all the findings in the Loukidelis order, and held that Canada's FPPs and their personal information practices are subject to the privacy law requirements in BC's PIPA, when they operate in British Columbia. PIPA applies to the FPPs because it applies to *all* organizations, including unincorporated associations. The judge situated this analysis within a broader endorsement of the significance of privacy for democratic values: "The ability of an individual to control their personal information is intimately connected to their individual autonomy, dignity and privacy. These fundamental values lie at the heart of democracy." All three FPPs have appealed the ruling to the BC Court of Appeals. At the time of writing, the litigation is ongoing.

Two further interesting developments signal the political and constitutional importance of this case. In April 2023, the government inserted in its budget legislation the brief provision (C.39) amending the CEA that states any political party may, subject to this act or any other applicable federal act, "collect, use, disclose, retain and dispose of personal information in accordance with the party's privacy policy" (Budget Implementation Act, 2023). This provision was widely regarded as a wholly cynical attempt to preempt the litigation in BC by signalling that there is indeed a regime for privacy protection at the federal level that ousts the provincial jurisdiction (Bennett, 2023). The government followed this up with a further amendment to the CEA (Bill C-65) introduced in March 2024. These actions provide the ammunition for the FPPs to collectively argue that Parliament has clearly intended to be the sole regulator of federal elections, including parties' data practices (Campbell, 2023a). The second development was the wholly unusual decision of the federal attorney general to intervene in the court of first jurisdiction in the BC case, regarded as an indication of the constitutional significance and sensitivity of the case (Campbell, 2023b).

The widespread implications should be clear from the strenuous and unified way the parties have fought the case. If they are subject to BC PIPA, then they would be required to amend their personal information practices to be in compliance with the BC law. Further, they would be forced to implement one set of standards in BC and another in the rest of the country. Beyond the constitutional arguments, this dispute signals the extent to which political parties are willing to fight for their individual, and collective, rights to process personal data without fear of intervention and oversight from Canada's privacy commissioners (Bennett, 2022).

We are left, then, with a profound paradox in the FPPs' positions. They want to argue for a consistent set of rules for the processing of personal data in campaigning across the entire country and object to one higher standard in BC. There is clearly a strong interest in a consistent set of rules and for a level-playing field. At the same time, they have steadfastly objected to the application of PIPEDA to political parties in their representations to the OPC. They have also refused to countenance any stronger provisions in the CEA beyond the minimal requirement of self-regulation through privacy policies, a "very timid step" according to the CCE (2021), and which contain no provisions for enforcement or independent oversight.

The "Closing of the Ranks" and the Collusion

Parties and politicians have long contended that their needs for personal information are special and that exempting political parties and representatives from privacy law supports freedom of political communication and enhances electoral and political processes (Cohen, 2021). But the initial reaction of FPPs to this pressure was largely one of silence, either out of a belief that the pressure would dissipate after a while or frankly out of a widespread ignorance of the scope of the problem.

The FPPs were first drawn out on the issue of privacy regulation at hearings before the House of Commons ETHI committee in 2018. The spokespersons for the main federal parties took different positions on the questions. For the Liberals, Mr. Michael Fenrick (2018) argued that "it would be a real disincentive to participation in the political process if people could face the kinds of penalties that exist for corporations, for instance, for non-compliance under PIPEDA." The Liberal Party of Canada did not support extending the application of PIPEDA in its current form to political parties, as "it's intended to address commercial activity. It's not intended to address political activity." The spokesman for the Conservatives, Trevor Bailey, conceded that their privacy policy was not in compliance with PIPEDA but declined to take a position on whether the law should apply to political parties. The NDP's director of operations, Jesse Calvert, stated unequivocally that the government should extend the application of PIPEDA to FPPs (ETHI, 2018, 21). Given more recent representations on these questions to the OPC, however, it appears that the NDP might have backtracked on this commitment; there was no mention of this pledge in either their 2019 or 2021 election platforms. Thus, the only FPP that is currently on the record as supporting the extension of federal privacy law to the FPPs is the Greens (Green Party of Canada, 2019).

Further evidence of resistance is evident from privacy law reform efforts. The government's efforts to reform PIPEDA with the Digital Charter Implementation Act (C-11) explicitly excluded any mention of political parties, contending that the very weak provisions in the Elections Modernization Act mandating voluntary privacy policies were sufficient (Mazereeuw, 2020). The same stance has been taken on the more recent iteration of this legislation (Bill C-27) in the face of constant calls by several witnesses before the House of Commons Industry Committee in Fall 2023. Perhaps more surprising, and troubling, was the reluctance of the Ontario government to follow BC's lead and include the regulation of provincial parties in its proposed reforms to Ontario's privacy law. An otherwise very progressive White Paper makes no mention of the regulation and oversight of Ontario's political parties (Ontario, 2021).

Efforts to strengthen the provisions in Quebec's Bill 64 relating to political parties, in line with prior recommendations by Élections Québec, were also met with strong resistance. It was reported that each of the main provincial parties caucused privately during the committee stage of the parliamentary process in order to develop a common position (Commission d'Acces, personal communication, October 12, 2022). The result was an insistence that any new privacy provisions be included in the Quebec Elections Act rather than in the newly amended privacy law (Quebec, Bill 64, Sections 84–90). The result is a set of provisions hedged by numerous exemptions and leaving doubt as to whether the new law significantly shifts the status quo with respect to Quebec's provincial political parties. And because the law only applies to bodies covered by the Quebec Elections Act, there is also little doubt that the law only applies to the provincial parties and not to the FPPs operating in the province.

In the BC litigation, we see very strong evidence that the FPPs have acted in concert to resist the commissioner's jurisdiction. In their submissions to the Office of the Information and Privacy Commissioner in response to the first complaint, each of the main FPPs filed separate briefs, and there were some differences between their arguments. In their petition to the Supreme Court of British Columbia (dated April 11, 2022) applying for judicial review of the Loukidelis decision, the lengthy submission submitted by the Liberal Party was joined almost entirely by both the Conservatives and the NDP. These parties also continued to support the Liberals' submission in the proceedings before the Supreme Court of BC and in the ongoing appeal of Judge Weatherill's decision of April 2024.

Further evidence of collusion is apparent from the reaction of the FPPs to the pressure to have their operations covered by the government's efforts to revise PIPEDA through its Consumer Privacy Protection Act. This bill is part of a larger package introduced in 2022 to give effect to the Digital Charter Implementation Act. The government's position has remained that political parties' personal information practices are properly regulated through the weak provisions within the Elections Modernization Act. There is now no daylight between the Liberals, Conservatives, or the NDP on the question of whether or not Canadian privacy protection law, overseen by the Office of the Privacy Commissioner of Canada, should apply to their operations. They are adamant that it should not.

Conclusion

The reform efforts and the litigation are continuing, and it is certainly premature to judge the final data protection rules for political parties that might emerge. At the moment, the argument that political parties are a unique breed of organization requiring special dispensations and exemptions from basic privacy standards seems to have held sway with both federal and provincial governments. I would make four general observations in conclusion.

First, and in a general and developmental sense, Canadian political parties exhibit increasing tendencies for technocratic, professionalized, and top-down campaigning, in contrast to grassroots mobilization. They embrace both digital campaigning practices and widespread voter analytics (Marland and Giasson, 2017). Personal data on supporters, members, and donors are jealously guarded. Analytical methods, such as the ideological scoring of voters, are regarded as confidential business practices to the extent that voters cannot legally discover their own scores, challenge them, or request their deletion – except, that is, in BC, where the commissioner explicitly rejected these arguments of commercial confidence and ordered the parties to disclose voter "scores" if they were requested (OIPCBC, 2019, 5).

Second, the close alliances between the FPPs and the consultancies and data mining companies that support their efforts are also threatened by the imposition of strong privacy regulations on political parties. Those companies do, of course, have to comply with the federal rules inherent in PIPEDA. To the extent, however, that political parties cannot rely on an unfettered access to data on the electorate, the analytical abilities of those companies are constrained. A good example would be Data Sciences, founded by Tom Pitfield, a friend of the prime minister and the spouse of the past-president of the Liberal Party. An exposé of Data Sciences in 2020 painted a picture of a powerful company, operating in the shadows, with extraordinary power to identify those voters most susceptible to influence (Castonguay, 2020).

Third, a simple rational-choice analysis of parties' incentives might conclude that those parties with the better data analytics operations, and the more sophisticated software, would be the parties most likely to resist regulation. Those behind in the race would have the stronger incentive to "level the playing field" through regulatory intervention. The initial positions of the parties before the ETHI committee in 2018 might have supported this thesis: the Liberals objected to privacy regulation, the Tories were on the fence, and the NDP were in support. Those distinctions are less apparent today. Indeed, the party that seems most strenuously opposed to the application of privacy law to both federal and provincial parties is the NDP.

Finally, the party cartel thesis relies on empirical evidence of collusion by parties to resist further regulation. Collusion is, by definition, a pattern of behaviour that will happen in secret and will be very difficult to observe empirically. The consistent pushback to the federal and BC commissioners would appear, on the surface, to reflect a level of collusion similar to that documented with respect to past resistance to ballot access and campaign financing rules (MacIvor, 1996). However, there does seem to have been a "closing of ranks" by the top three Canadian federal parties, at least for now. Whether or not this represents a party "cartel," however, will depend on long-term evidence of the consolidation of data resources within the top three political parties and a progressive exclusion of smaller players who have neither the resources nor the technical expertise to wage an effective modern digital campaign.

Notes

[1] The research for this chapter is funded by a SSHRC Insight Grant on "Micro-Targeting and Data-Driven Elections in Canada," 435-2019-0403.

References

Bennett, C.J. (2013). The politics of privacy and the privacy of politics: Parties, elections and voter surveillance in Western democracies. *First Monday*, *18*(8). https://firstmonday.org/ojs/index.php/fm/article/view/4789.

Bennett, C.J. (2015). Trends in voter surveillance in Western societies: Privacy intrusions and democratic implications. *Surveillance and Society*, *13*(3–4).

Bennett, C.J. (2018). Data-driven elections and political parties in Canada: Privacy implications, privacy policies and privacy obligations. *Canadian Journal of Law and Technology*, *16*(2).

Bennett, C.J. (2020). Federal Court File No. T-893-20; Bennett Expert Opinion Report in CCE v. CDR.

Bennett, C.J. (2022, March 23). BC privacy ruling over political party data collection is a victory for voters' privacy. *The Globe and Mail.*

Bennett, C.J. (2023, May 1). Government's efforts to introduce privacy rules for federal political parties wholly inadequate and totally cynical. *The Hill Times.* https://www.hilltimes.com/story/2023/05/01/governments-efforts-to-introduce-privacy-rules-for-federal-political-parties-wholly-inadequate-and-totally-cynical/385273/.

Bennett, C.J., and Bayley, R.M. (2012). *Canadian federal political parties and personal privacy protection: A comparative analysis*. Office of the Privacy Commissioner of Canada.

Bennett, C.J., and Gordon, J. (2021). Unpacking the micro in micro-targeting: An analysis of Facebook advertising in 2019 federal election. *Canadian Journal of Communications*, *46*(3).

Bennett, C.J., and McDonald, M. (2020). From the doorstep to the database: Political parties, campaigns, and personal privacy protection. In N. Witzleb, M. Paterson, and J. Richardson (Eds.), *Big data, political campaigning and the law* (141–63). Routedge.

Bennett, C.J., and Oduro Marfo, S. (2019). Privacy, voter surveillance and democratic engagement: Challenges for data protection authorities. *International Conference of Data Protection and Privacy Commissioners*.

British Columbia Supreme Court. (2022). *Federal Political Parties v. the Complainants* S-223033; S-223065; S-223104.

British Columbia Supreme Court. (2024). Liberal Party of Canada v. the Complainants. BCSC 814 (CanLII). https://canlii.ca/t/k4lk3.

Burkell, J., and Regan, P.M. (2019). Voter preferences, voter manipulation, voter analytics: Policy options for less surveillance and more autonomy. *Internet Policy Review*, *8*(4), 1–24.

Cadwalladr, C. (2017, May 7). The great British Brexit robbery: How our democracy was hijacked. *The Observer*. https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy.

Campbell, I. (2023a, September 11). Federal parties look poised to use new law to fight B.C. privacy regulations, court documents indicate. *The Hill Times*. https://www.hilltimes.com/story/2023/09/08/federal-parties-appear-poised-to-use-recent-legislation-to-fight-against-b-c-privacy-laws-court-documents-indicate/396723/.

Campbell, I. (2023b, December 11). Attorney General makes "highly unusual" decision to get involved early in privacy case involving federal political parties. *The Hill Times*. https://www.hilltimes.com/story/2023/12/11/attorney-general-makes-highly-unusual-decision-to-get-involved-early-in-privacy-case-involving-federal-political-parties/405761/.

Canada Bill C-47. (2023). *Budget Implementation Act, 1st Sess. 44th Parl, 2023*. Division 39.

Canada Bill C-76. (2018). *Elections Modernization Act: An act to amend the Canada Elections Act and consequential amendments, 1st Sess, 42nd Parl, 2018*.

Castonguay, A. (2020, January 8). La nouvelle machine a gagner des elections. *L'actualite*. https://lactualite.com/politique/la-nouvelle-machine-a-gagner-des-elections.

Centre for Digital Rights. (nd). https://www.centrefordigitalrights.org/about.

Cohen, T. (2021). The political exemption: A justifiable invasion of privacy in the political sphere?. *The University of New South Wales Law Journal*, *44*(2), 584–612.

Commissioner of Canada Elections. (2021). *Annual report 2021*. https://www.cef-cce.ca/content.asp?section=rep and dir=ar/rep10 and document=p1 and lang=e#b0c6.

Council of Europe. (2021). *Guidelines on the protection of individuals with regard to the processing of personal data by and for political campaigns*. https://rm.coe.int/guidelines-on-data-proetction-and-election-campaigns-en/1680a5ae72.

Curry, B. (2019, June 13). Majority of poll respondents express support for extending privacy laws to political parties. *The Globe and Mail*. https://www.theglobeandmail.com/politics/article-majority-of-poll-respondents-express-support-for-extending-privacy/.

Dalton, R.J., and Wattenberg, M.P. (Eds.). (2002). *Parties without partisans: Political change in advanced industrial democracies*. Oxford University Press.

Delacourt, S. (2016). *Shopping for votes: How politicians choose us and we choose them*. D and M Publishers.

Dubois, E., and Taylor, O. (2019). *Understanding the digital ecosystem: Findings from the 2019 federal election*. The Digital Ecosystem Research Challenge Report. Pol Comm Tech Lab. https://www.polcommtech.com/our-research/derc-report.

Ekos Research Associates. (2020). Survey of candidates of the 43rd general election: Narrative report. https://epe.bac-lac.gc.ca/100/200/301/pwgsc-tpsgc/por-ef/elections_canada/2020/2020-05-e/Narrative_Report-EN.pdf.

Elections Canada. (2013). *Preventing deceptive communications with electors: Recommendations from the Chief Electoral Officer of Canada following the 41st General Election*.

Élections Québec. (2019). *Partis politiques et protection des renseignements personnels: exposé de la situation québécoise, perspectives comparées et recommandations.*

Fenrick, M. (2018). *Evidence - ETHI (42-1) - No. 123*. House of Commons of Canada.

Gauja, A. (2014). Building competition and breaking cartels? The legislative and judicial regulation of political parties in common law democracies. *International Political Science Review*, *35*(3), 339–54.

Government of Ontario. (2021). *Modernizing privacy in Ontario: Empowering Ontarians and enabling the digital economy*. White paper.

Green Party of Canada. (2019, September 17). *Securing your data: Green Party makes privacy protection a priority* [Press release].

House of Commons, Standing committee on Access to Information, Privacy and Ethics. (2018). *Addressing digital privacy vulnerabilities and potential threats to Canada's democratic electoral process: Report of the Standing committee on Access to Information, Privacy and Ethics.*

Howard, P.N., and Kreiss, D. (2010). Political parties and voter privacy: Australia, Canada, the United Kingdom, and United States in comparative perspective. *First Monday, 15*(12).

Issenberg, S. (2013). *The victory lab: The secret science of winning campaigns*. Random House.

Judge, E.F., and Pal, M. (2021). Voter privacy and big-data elections. *Osgoode Hall LJ*, *58*, 1–55.

Katz, R.S., and Mair, P. (1995). Changing models of party organization and party democracy: The emergence of the cartel party. *Party politics*, *1*(1), 5–28.

Katz, R.S., and Mair, P. (2009). The cartel party thesis: A restatement. *Perspectives on Politics*, *7*(4), 753–66.

Koops, B.J. (2021). The concept of function creep. *Law, Innovation and Technology*, *13*(1), 29–56.

Leblanc, D., and Cardoso. T. (2019. April 24). PMO vets potential judges with private Liberal database. *Globe and Mail*. https://www.theglobeandmail.com/politics/article-pmo-vets-potential-judges-with-liberal-database/.

MacIvor, H. (1996). Do Canadian political parties form a cartel? *Canadian Journal of Political Science / Revue canadienne de science politique*, *29*(2), 317–33.

Marland, A., and Giasson, T. (2017). From brokerage to boutique politics: Political marketing and the changing nature of party politics in Canada. In A.G. Gagnon and A.B. Tanguay (Eds.), *Canadian parties in transition* (343–63). University of Toronto Press.

Mazereeuw, P. (2020, November 21). "Massive overhaul" of privacy law leaves political parties off the hook. *The Hill Times*. https://www.hilltimes.com/story/2020/11/21/massive-overhaul-of-privacy-law-leaves-political-parties-off-the-hook/228868/.

Newman, K. (Host). (2019, September 26). Attention Control: The data on us. Podcast with anonymous whistleblower. https://www.listennotes.com/podcasts/attention-control/the-data-on-us-GTyS6vKpWIF.

Office of the Information and Privacy Commissioner of British Columbia. (2011). *Summary of the Office of the Information and Privacy Commissioner's investigation of the BC NDP's use of social media and passwords to evaluate candidates*.

Office of the Information and Privacy Commissioner of British Columbia. (2013). *Sharing of personal information as part of the draft multicultural strategic outreach plan: Government of British Columbia and BC Liberal Party.* Investigation Report F13-04.

Office of the Information and Privacy Commissioner of British Columbia. (2019a). *Full disclosure: Political parties, campaign data and voter consent*. Investigation Report P19-01.

Office of the Information and Privacy Commissioner of British Columbia. (2019b). *Order P19-02. Courtney-Alberni Riding Association of the New Democratic Party of Canada*.

Office of the Information and Privacy Commissioner of British Columbia. (2022). *Order P22-02. David Loukidelis QC. Conservative Party of Canada, Green Party of Canada, Liberal Party of Canada, New Democratic Party of Canada.*

Office of the Privacy Commissioner of Canada. (2018, September 11–13). *Securing trust and privacy in Canada's electoral process*. Resolution.

Office of the Privacy Commissioner of Canada. (2019, April 1). *Guidance for federal political parties on protecting personal information*.

Office of the Privacy Commissioner of Canada. (2021, March 25). *Letter regarding complaint against federal political parties*.

Quebec National Assembly. (2021). *Bill 64, An Act to modernize legislative provisions as regards the protection of personal information.*

Rubinstein, I. (2014). Voter privacy in the age of big data. *Wisconsin Law Review*, *5*, 861–936.

Scassa, T. (2018, May 2). A federal bill to impose privacy obligations on political parties in Canada falls (way) short of the mark. www.teresascassa.ca.

Scassa, T. (2019). Decision paves the way for federal riding associations in BC to be subject to BC's data protection laws. www.teresascassa.ca.

Therrien, D. (2018). *Appearance before Standing Committee on Procedure and House Affairs on the study about Bill C-76*. Elections Modernization Act. Office of the Privacy Commissioner of Canada.

United Kingdom Information Commissioner's Office. (2018). *Democracy disrupted: Personal information and political influence*.

United Kingdom Information Commissioner's Office. (2019). *Investigation into the use of data analytics in political campaigns. A report to Parliament*.

United States Federal Trade Commission. (2019, July 24). *FTC imposes $5 billion penalty and sweeping new privacy restrictions on Facebook*.

van Biezen, I., and Kopecký, P. (2014). The cartel party and the state: Party-state linkages in European democracies. *Party Politics*, *20*(2), 170–82.

Wylie, C. (2019). *Mindf*ck: Cambridge Analytica and the plot to break America*. Random House.

Young, L. (1998). Party, state and political competition in Canada: The cartel model reconsidered. *Canadian Journal of Political Science / Revue canadienne de science politique*, *31*(2), 339–58.

5

Digital Data as a Lens on Voters' Lifestyle

Theoretical Perspectives and Insights

Catherine Ouellet and Yannick Dufresne

Introduction

The craze for big data is undeniable. Over the last few years, the possibilities for retrieving and storing these data in various forms have multiplied. Operations that once required a supercomputer can now be performed on a smartphone. There is virtually no limit anymore to the amount of information that can be processed using computers and AI. Unsurprisingly, politicians have been quick to take advantage of these new opportunities in the way they do politics. One notable example is the recent use of the Liberal Party's database by the Trudeau Government to conduct systematic checks on candidates for judicial appointments – a practice abandoned in the face of widespread criticism from both political opponents and legal experts (Leblanc, 2021). Social media and data-driven companies such as Google, Twitter, and Facebook have also played an increasingly influential role in recent campaigns, where the billions of data generated allow political parties to refine their micro-targeting techniques in unprecedented ways (see Judge and Pal [2021] for a critical overview of these techniques in the Canadian context). Big data has therefore become indispensable for political parties' quest for power, playing a pivotal role in electoral competition. By harnessing vast quantities of digital information, parties are now able to tailor their strategies and messages with unprecedented precision, targeting potential voters more effectively than ever before (Delacourt, 2013; Flanagan, 2009). Google sells granular user data more than 19.6 times per minute to different companies (Gendreau, 2023). Not to mention Amazon, Apple, Microsoft, and the other web giants. The volume of data that we create – and which is used – every day is mind-boggling. The widespread use of big data and data analytics by both political parties and the private sector has serious implications for the health of our democracy – an issue to which we will come back later.

While this new reality has changed politics, it has also impacted the way social scientists conduct research. For instance, studies using complex digital data still face several challenges related to the central characteristics that define big data. These are commonly described as the "three V's": volume, variety, and velocity (Laney, 2001), with now a quite large definitional consensus around them (Emmanuel and Stanier, 2016).

Volume refers to the massive amounts of data or the magnitude of data generated from different sources – such as cell phones or social media. Velocity refers to the speed at which these amounts of data are received, stored, processed, and analyzed. Variety emphasizes the idea that we no longer deal with structured data only; these can take various forms, and most current data is mostly semi-structured or unstructured – such as images, audio, or social media updates (Emmanuel and Stanier, 2016).

Due to their characteristics, the ability to collect and analyze the new data introduces technical, hosting, and accessibility issues but also difficulties in terms of sharing and understanding the data. Concurrently, the digitization of privacy inevitably raises a certain number of ethical issues (see Déziel, this volume). Some citizens are even concerned that massive data and technology will take precedence over humanity (Lazer and Radford, 2017; Zuboff, 2019). We can also think of successful films such as *Ex Machina* by Alex Garland (2014), *Terminator* by James Cameron (1984), or *Her* by Spike Jonze (2013), which question the possible superiority of artificial intelligence over humans (see Wilson, 2018). The highly-mediated scandal of Cambridge Analytica and Facebook already showed how the "datafication"[1] of the world also raises data privacy and security issues – but more generally, highlights the potential harmful effects of AI and big data for electoral democracies. Besides fraudulent practices, critics are also concerned by – among other things – an increasing polarization, the spread of disinformation, and the approach of electors as consumers (see Judge and Pal, 2021).

But if the entry into the digital age comes with its share of legitimate criticism, it also paves the way for innovative research opportunities. In this context, we need, as social scientists, to ask ourselves big questions. How can these vast volumes of data be made accessible to further science and contribute to a better understanding of the world? What are the potential contributions of new digital data to our understanding of social phenomena? Conversely, does exposure to these new data have an effect on individual behaviour? And what are the risks, if any? To what extent are these changes potentially harmful for the vitality of electoral democracies? Therefore, researchers need to think both about the validity and the interpretation of these digital traces to understand the social world. At the same time, critical thinking is a worthwhile, essential skill for any data scientists as they constantly need to ask important questions but also question the data they gather and work with. It is in this perspective that this chapter highlights the main promises and challenges of the use of massive digital data by social scientists and practitioners. More particularly, it also examines the theoretical and empirical potential of lifestyle data – which often take the form of digital traces – on our understanding of Canadian politics. It examines what these data reveal about the political behaviour of voters, a crucial question that is still underexplored in the social sciences but also a prime example of how the digitization of our lives allows us, as social scientists, theoretical advances. In parallel, it shows the degree of surveillance to which we, as citizens, are subjected (see also Lyon, this volume). Finally, this chapter aims to provoke a reflection on the value and the potential risks of lifestyle data in a digital world.

What Can We Learn from Digital Data?

While the digitization of personal and public data raises a certain number of important questions, it nevertheless represents a phenomenon of major interest for the research community. The use of data, its interpretation, and what it ultimately has to tell us about the world opens the door to promising theoretical and empirical advances.

In recent years, several researchers have demonstrated the extent of the theoretical advances made possible by this explosion of new data. For example, digital data from social media allow us to study traditionally underrepresented populations.[2] Twitter has been a tool in the study of the behaviour of people with mental health problems (Coppersmith, Dredze, and Harman, 2014; De Choudhury et al., 2016) and has made it possible to identify citizens who played a pivotal role in the emergence of social movements such as Black Lives Matter or Occupy Wall Street (Barberá et al., 2015; Jackson and Foucault Welles, 2016). In this regard, social media are becoming essential gateways for studying citizen behaviour and the logic of social movements (Lazer and Radford, 2017). These new data also allow for the study of the impact of specific political issues on the behaviour and attitudes of subgroups of voters. For example, to what extent did the Conservative Party of Canada's (CPC) positioning on the issue of Israel in 2011 influence the voting behaviour of Canadian Jews? The use of massive digital data thus makes it possible to fragment the electorate into micropublics and to examine, for instance, the impact of public policies on specific communities.

New digital data also allows for the study of various phenomena with an impressive granular temporal and geographical precision (Lazer and Radford, 2017). In this regard, through billions of Twitter metadata, Peter S. and colleagues (2011) have access to signals coming directly from a vast number of individuals over time and over space, revealing the potential of Twitter to describe universal human patterns such as happiness. Through a dataset comprising 4.6 billion expressions posted over a thirty-three-month span by over 63 million unique users, the researchers uncovered and explained temporal variations in happiness and information levels over timescales with unprecedented granularity. Similarly, researchers have examined interaction networks involving millions of people on Twitter and Yahoo to make electoral predictions that would have been unthinkable a few years ago (Beauchamp, 2017; State et al., 2015). These studies reveal deep "intra-cultural correlations" between individuals spread across the globe – and thus opportunities for studying human systems as composed of subsystems and individuals dynamically connected according to their "ways of being" and geospatial location (Lazer and Radford, 2017, 25).

Other researchers have used web search engines to show how these can be used to accurately track outcomes such as unemployment levels (Askitas and Zimmermann, 2009), auto and home sales (Choi and Varian, 2009a, 2009b), or even disease prevalence in near real time (Polgreen et al., 2008). In this vein, Goel and colleagues (2010) show that what consumers are searching for online – such as attending movies and purchasing music or video games – is predictive of their collective future behaviour days or weeks in advance. These data are all *digital traces*, or records that chronicle actions taken. Other examples of this type of big data include governmental data such as voter records (Bonica, 2014), tax data (Chetty, Hendren, and Katz, 2016), mobile communication networks data, the last one being used for studying the tie strengths between individuals (Onnela et al., 2007), or even to predict unemployment rates ahead of official reports and more accurately than traditional forecasts (Toole et al., 2015).

Other studies rely on digitalized life data per se, which is more than the record of an action but the action itself (Lazer and Radford, 2017). Here, we can think of Michel and colleagues (2011), who constructed a corpus of digitized texts to take a novel approach to the study of culture. With a corpus containing around 4 percent of all books ever printed, the authors explore linguistic and cultural phenomena that were reflected in the English language between 1800 and 2000. Other examples in this regard include scanning newspapers to study the dynamics of fame (van de Rijt et al., 2013) or using digital records of more than two thousand newspapers and online news sources to explain the persistent underrepresentation of women in news coverage, thereby offering an important contribution to the sociology of the media and the sociology of gender (Shor et al., 2015).

These works are obviously far from exhaustive of all the new research avenues offered by the digitalization of life. But these few examples have at least one thing in common: they show how massive digital data can help us understand previously hidden as well as novel microscale mechanisms underlying the social world.

A Theoretical Challenge

If the digital era indeed opens up new avenues of research (as we have just seen with a couple of fruitful examples), the quantification of human life by digital information can also provide a sea of useless numbers if it is detached from any theoretical thinking. Therefore, it would be wrong to think that the use of massive and digital data makes theory building easy, or easier than a few decades ago. On the contrary, researchers must act as a counterweight against the rise of "big data hubris" – or the belief that the volume of big data can substitute (rather than supplement) traditional data and allow us to draw conclusions from numbers alone (Lazer et al., 2014). More generally, the social sciences must counter a blind enthusiasm for the datafication of social explanation.

As we have seen, the digitization of our lives and the creation of new spaces are influencing the social world itself and are causing new social phenomena – for example, the effect of Facebook or Twitter on selective information exposure (Cinelli et al., 2020) or the consequences of Trump's use of social media on international relations (see, for example, Duncombe, 2019; Zeitzoff, 2017). Therefore, it goes without saying that these questions must be studied within the framework of the social sciences. But now that massive digital data are at our fingertips and are accumulating every day, social researchers have to know what to do with them. As recently stated by Lazer and Radford (2017), in the coming years, the question "will be less about whether the data exists, but more about what we can study with the available data" (28). To put it otherwise, we need to circumscribe the theoretical scope of these new data. Without a theoretical postulate and substantial interpretation of the data, these are nothing but numbers.

The next section explores the concept of lifestyle and what it can bring to political science research. As lifestyle characteristics often take the form of digital traces – we produce these traces daily, for instance when we share our favourite music on Spotify or other open databases, when we shop on Amazon, or even when we order food online – the concept of lifestyle provides an opportunity to dig deeper on how new digital data can help us better understand the social world. The concept of lifestyle is coming from marketing research,[3] but it is still largely unexplored in political science research.

The digital paradigm shift is now permeating the contemporary political world. For instance, massive digital datasets are now central to most election campaigns and is used to target specific subgroups of voters. Barack Obama's 2008 presidential campaign, whose team used public data to personally target voters, was the first historical example of branding (Nickerson and Rogers, 2014). Canadian political parties, although less resourced, developed similar political strategies in the following years (Delacourt, 2013; Dufresne and Marland, 2012; Flanagan, 2009). Web professionals are no longer on the periphery, they are actively engaging in web campaigns and tailoring political messages. Electoral politics increasingly hinge on data insights, with more and more traditional volunteer efforts being replaced with specialized expertise (see Biancalana, this volume). The Conservative team even collaborated with former Australian and American political strategists to import their more advanced political marketing practices. At the Canadian federal level, the CPC was the first to rely on a sophisticated private database for electoral purposes – a database built in 2004 under Stephen Harper and named the Constituent Information Management System (CIMS). In 2006, when Stephen Harper first became prime minister, the CPC's strategists were relying on CIMS to practice an ongoing analysis of the Canadian electorate based on lifestyle characteristics: "A big part of [their] job was to consider everything the Conservatives did in terms of potential gains or losses in future elections, twenty-four hours a day, seven days a week, similar to how Canadians were consuming their products and services" (Delacourt, 2013, 204). During the 2000s, the Conservative team exemplified people who would never vote for them with Zoé, a "twenty-five [year-old] single [girl], with a degree in sociology, practicing yoga and eating organic food" (Flanagan, 2009, 223–24). Twenty years later, the use of lifestyle data is now widespread among Canadian political parties, which makes this concept even more relevant to illustrate some important challenges and questions raised by the digital era.

Lifestyle and Politics

One of the most acknowledged stabilizing factors of party systems has been the strong impact of social characteristics on voting. From the mid-1940s, it has been established that "a person thinks, politically, as he is, socially. Social characteristics determine political preferences" (Lazarsfeld et al., 1944, 27). The relationship between social cleavages – such as religion, economic status, or location – and party systems has been a major question for political comparativists during the following decades (Lipset and Rokkan, 1967; Rose, 1974; Franklin 1992). It is now pretty consensual that for a long time, these historically rooted societal differences have shaped groups of individuals with common interests, which in turn have been reflected in party loyalties and electoral support. However, in recent decades, the relevance of the cleavage-party nexus has been revisited, with many scholars questioning the once robust effects of social structure on political cleavages (Dalton, 1984; Franklin, 1992). In face of the weakening of traditional social identities, some of them concluded that voters' choices in Western democracies were becoming less predictable on a cleavage basis (Dalton, 1984, 2000; Franklin, 1992). In this context, Clark and Lipset (1991) concluded that politics in postindustrial societies was now "less organized by class and more by other loyalties" – such as gender, ethnicity, or consumption sectors (Huber and Suryanarayan, 2016; Norris and Mueller, 1988; Saunders, 1981, 1990). This also implied a more volatile electorate, suggesting that more importance was now given to short-term factors such as issues, performances judgments, or candidate image in the balance of vote choice – some suggesting the development of "an eclectic and egocentric pattern of political decision making" across democracies (Dalton and Wattenberg, 1993, 213). Judging from this evidence, long-term influences such as socialization, or personality traits, would only play a weak role in explaining contemporary politics.

These transformations took place in a context of social and cultural fragmentation. Indeed, the increasing prosperity that accompanies modernity (Giddens, 1991) leads to an increased individualization, where individuals behave from now on according to their personal values and their particular preferences. As explained by Bishop (2008), "We were witnessing something more fundamental – a cultural shift fueled by prosperity and financial security. Comfortable and carefree, people are reorganizing their lives according to their values, tastes, and beliefs." In this context, the notion of choice takes on a whole new dimension: modernity confronts the individual with a multitude of choices. For sociologist Anthony Giddens (1991), "One of them concerns the preponderance of lifestyle, and its inevitability for the individual agent. Under the conditions of post-modernity, the individual is not only free to follow a lifestyle – but has no choice but to choose it." We therefore see a fragmentation of traditional practices, coupled with an increasing diversification of individual interests and preferences (Fischer and Mattson, 2009; Lamont et al., 1996). This leads to a growing heterogeneity of individual behaviours within social classes: the relationship between social class and behaviour is eroded (Dalton, 2000; Franklin, 1992). Consequently, researchers have come to talk about a "pluralization of lifestyles" (Berger, 1973). For instance, some researchers have also suggested that the number of "distinct social worlds" or "cultural enclaves" in the US have multiplied between 1970 and 2005 (Fisher and Mattson, 2009; see also Florida, 2008).

This decrease in the capacity of traditional sociological models to predict voting behaviour has led some researchers to focus on "lifestyle politics," the idea that individuals will engage in a certain form of lifestyle to tackle environmental, ethical, or other societal issues (Bennett, 1998; de Moor, 2017; de Moor and Verhaegen, 2020; Micheletti and Stolle, 2004). Indeed, the fragmentation of socialization units – style of music, aesthetic or artistic allegiance, daily lifestyle, style of dress, eating habits, and so on – could help explain individual values, preferences, attitudes, and opinions (Chaney, 2002; Stolle, Hooghe, and Micheletti, 2005; Weiss, 1988). And among the few empirical studies on the subject (mostly in the US), scholars demonstrate that even our smallest choices, such as coffee preferences, speak volumes about us. The work of DellaPosta, Shi, and Macy (2015) and Mutz and Rao (2018) on the reasons behind American progressives' attachment to lattes are an example. From a psychological perspective, the recent book *Prius or Pickup?* by Marc Hetherington and Jonathan Weiler (2018), which explains how the food we eat, the music we listen to, or the car we drive reveals so much about the American red/blue divide, is another prime example.

These relationships, while they may seem trivial at first, are extremely important in that they have profound implications for how we should understand our political life. And their strength also suggests that the individualization thesis may be overstated (see Rose and McAllister, 1986). Indeed, the weakening of traditional social identities on vote choice does not imply that social groups are not relevant to understanding political preferences. It may only be the case that in a context of deep cultural and social fragmentation, the boundaries of these social groups need to be redefined – social groups that lifestyle variables might be able to delineate and capture.[4] These fragmentary works therefore reassess the importance of socialization on the formation of political preferences and contribute to the growing body of research suggesting that political differences extend far beyond overtly ideological opinions to much subtler and more banal preferences, tastes, and lifestyles. Besides, if uncovering the underlying essence of the vote has been a central endeavour of electoral behaviour studies over the past seventy years, it is also a matter of practical interest. For instance, when sociological factors explain voters' electoral preferences, very little room is left for individuals to independently determine their issue preferences. Indeed, a sociological explanation is hardly consistent with some of the fundamental assumptions of political marketing, as the marketing logic involves targeted micromessages to specific publics, thus implying an issue-based and volatile voting behaviour conception (Plasser and Plasser, 2002). Today, politicians are speaking in shopping language. According to Patrick Muttard, top strategist for the Conservative team between 2004 and 2006, trying to expand Conservatives' support would mean "going to Tim Hortons, not to Starbucks" (Delacourt, 2013, 11). Research at the intersection of lifestyle and politics therefore assesses the extent to which these strategies are effective.

Studying Lifestyle Differently

The decreasing predictive power of traditional sociodemographic variables on vote choice leads us to rethink the way public opinion research is conducted. Indeed, if contemporary quantitative studies in political science have, for a long time, made use of public survey data from public and electoral polls, there is now a need to broaden and deepen these databases by incorporating diverse variables – lifestyle characteristics being some of them – but also diverse data sources such as from social media or open digital databases. In this regard, DellaPosta and colleagues

(2015) depict very well how, in this specific case, traditional survey data are constraining researchers from making theoretical advances in the field of lifestyle and politics:

> We then pursue possible solutions to the puzzle of lifestyle politics, beginning with a review and critique of the prevailing explanatory strategy ... These explanations are intuitively compelling and often well documented with data from opinion surveys, but they are constrained by the absence of network relations in the random samples on which most surveys are based. Decades of survey-based opinion research have lacked the relational data with which to rule out network autocorrelation as an alternative explanation. (1476)

Massive digital data is one important pillar of this research methodologies' reorientation. For instance, let's stick with DellaPosta and colleagues (2015), who postulate that the relationship between lifestyle and electoral behaviour is rooted in social influence and homophily processes. For them, digital data represents a unique opportunity to study public opinion dynamics embedded in networks of interactions (1503). Given the worldwide tendency for individuals to interact using digital devices that record their interactions, digital data therefore become a gold mine for these researchers. Taking the example of some studies using social media data to measure the impact of homophily on social interactions (see Aral and Walker, 2012; Bond et al., 2012), the authors also stress that "these early studies inspire [them] to close on a note of optimism, by pointing to the intriguing possibility that social science is poised for a new relational beginning" (DellaPosta, Shi, and Macy, 2015, 1503). Similarly, to measure the potential predictive and explicative power of lifestyle variables on political preferences, we need to think beyond traditional survey methods and collect data that allow for more granular but also dynamic analyses.

Besides, as stressed earlier, the widespread use of lifestyle-related data for political purposes is now a well-known fact. Canadian political strategists – who are more likely to have a degree in mathematics or computer science than in political science – rely on parties' databases, which contain granular information about voters, to refine their micro-targeting techniques and tailor their messages (Delacourt, 2013). As social scientists interested in lifestyle, not only do we want to take advantage of the theoretical potential of this rich, meaningful concept in the discipline, but we also want (fortunately) to understand how lifestyle data – and big data more broadly – are used by political parties and, more importantly, to be at the forefront of their political marketing techniques.

Exploring the Canadian and the Quebec Cases through Datagotchi

The migration of modern life to new social structures opens up opportunities to observe and measure human behaviour as never before. As shown earlier, new digital data allow us to study, from a more global perspective, the shifting complexity of the world while at the same time focusing on individual human behaviour. Digital data also open the door to new ways of measuring public opinion.

In this context, we codeveloped[5] an online, educative, and playful data collection tool called Datagotchi,[6] which predicts individual vote choice based on lifestyle characteristics. By answering different questions about their lifestyle (leisure activities, material and cultural preferences), users personalize their avatars and their environment. At the end, based on the answers given during the questionnaire, Datagotchi predicts users' probability of voting for each of the main parties in a given election. The application was launched for the first time during the 2021 federal elections in Canada and already received extensive media coverage – the application received over 350,000 visits and more than 30,000 completed questionnaires for data analysis. In partnership with the Canadian Press, Datagotchi was launched again during the 2022 general elections in Quebec, collecting over 90,000 completed questionnaires.[7] The objectives of Datagotchi are threefold: (i) creating a scientific, playful, and secure research tool to better understand myriad phenomenon – such as the relationship between lifestyle and voting behaviours; (ii) through interactive and popularized analyses, helping citizens to better understand more complex data models; and (iii) make citizens aware of the wealth of individual data produced by each of us every day and which often take the form of digital traces. In this perspective, Datagotchi is intended to both take advantage of massive digital data to our understanding of human behaviour and to be a sort of awareness-raising tool in the face of large-scale use of lifestyle data (and other personal data) for either political or marketing purposes.[8]

Generalizing empirical research usually calls for a probability-based random sample, seen as the gold standard. Data from opt-in surveys like Datagotchi, however, pose unique challenges as they diverge from the traditional random sampling methods elucidated in classic survey theory (Couper, 2017). While Datagotchi skews towards younger participants with left-leaning political orientations and has a marginal overrepresentation of women and more educated individuals, contemporary methodologies provide ways to mitigate these biases (see Couper, 2017; Wang et al., 2015). Notably, statistical techniques such as propensity score matching or reweighting allow researchers to account for nonrandom selection in observational datasets. When properly applied, these methods can help achieve balance across observed characteristics, making the dataset more similar to a randomized experiment.

Rather than using a conventional reward-based model where participants are incentivized through monetary rewards or points, Datagotchi relies on the appeal of game dynamics. Given the pervasive survey fatigue (Van Mol, 2017), where individuals are inundated with monotonous questionnaires, a gamified model, while potentially less representative, can score exceptionally high on data validity (see also Peer et al., 2022). Finally, incorporating advanced expertise and methodological rigour from the academic community can potentially enhance the business model of professional survey firms, particularly in their adoption of web panels for data collection. While polling firms vary in their techniques (e.g., in the choice of sample, the way they calculate margins of error, or the way they ask questions), the scientific research community typically follows a more systematic approach. This academic rigour is further strengthened by the oversight of ethics board committees, which impose stringent guidelines to ensure the integrity and ethical soundness of research practices (see Page and Nyeboer, 2017). The academic model also gives researchers complete control over the data collection process and allows them to guarantee the quality of the collected data.

Exploring Musical and Movie Preferences

Datagotchi explores various dimensions of lifestyle including, among others, musical and movie preferences. Users are asked to report their favourite musical artist or group as well as their favourite film of all time. Using this information, Datagotchi allows researchers to automatically collect from online databases[9] additional information regarding the indicated artist/group and movies (music or movie genre, year of production, etc.), thus complementing the lifestyle profiles of users. From there, we are – among other things – able to look at how Quebecers vary in their political attitudes based on their favourite musical genre. For instance, empirical analysis suggests that Conservative voters have a disproportionate tendency to prefer metal or country music compared to other groups of voters, while folk music is more popular among left-wing party voters (see Ouellet and Fréchet, forthcoming). Interestingly, these results remain consistent across elections but also with previous work conducted in the United States (see Ouellet and Fréchet, forthcoming). And the predictive power of musical preferences on vote choice is not that surprising, since music preferences are often reflective of personality traits or other aspects of our identities (see MacDonald, Hargreaves, and Miell, 2017; Rentfrow, McDonald, and Oldmeadow, 2009). So far, music preferences have been shown to be correlated with sociodemographic markers such as class, age, or gender (Rentfrow and Gosling, 2007) but also with personality traits (Litle and Zuckerman, 1986; Zweigenhaft, 2008) and other lifestyle-related activities (North and Hargreaves, 2007). The collected data through MusicBrainz and OMDb also allows for more sophisticated analysis, such as multidimensional scaling, mappings between sociodemographic attributes or political attributes, and musical preferences (see Goldberg, 2011). Data from Datagotchi also allow for dynamic analyses, such as the creation and monitoring of voter segments akin to those utilized by political parties, tracking their progression throughout a campaign. Furthermore, it provides a means to gauge the effects of campaign events like surveys or televised debates on specific subsets of the electorate.

Ethical Issues and Conclusion

Although preliminary, these analyses shed light on some important questions raised by the production (and use) of data in the digital age. While, nowadays, it is common for people to share their favourite playlists through various music streaming services, musical preferences are rarely considered sensitive information. Conversely, this information is often considered rather trivial. However, this kind of information is collected and analyzed and allows streaming platform algorithms to make increasingly reliable suggestions in line with their consumers' preferences. Daily, these digital traces therefore allow new insights into users' preferences. And as social scientists, this information is finally not that trivial as it allows us to discriminate between social groups and to map out groups of voters (see Ouellet and Nadjim, forthcoming). We therefore need to engage in a broader ethical reflection on the digital traces that we continuously produce and, despite their great value understanding of the social world, be continually aware of the challenges and issues they raise – especially since they are, in certain contexts, within the reach of political strategists for political purposes.[10] The development of innovative data collection tools such as Datagotchi also involves its share of ethical considerations, as the confidentiality and security of the data collected must never be compromised.

This phenomenon inevitably raises larger questions. For example, what are the rights of the secondary subjects who produce – often without consent – these data? What is the scientific value of documents obtained by hackers like Wikileaks and the Panama Papers? How can we reconcile the norms of transparency and reproducibility of research with the inevitable loss of confidentiality caused by the current technical capacities to identify individuals through the crossing of databases? What is the evolution of the social acceptability of this loss of confidentiality and what is its social impact? These are legitimate and important questions. And as social scientists, we constantly need to reflect on the security and the confidentiality of the data we use.

As we know, the exploitation of such data by political parties is now the norm. With the help of strategists and analysts, political parties are using increasingly sophisticated data-driven methods to dissect and target specific segments of the electorate. Some recent events illustrate the ethical issues raised by such practices – the scandal related to Facebook and the firm Cambridge Analytica being a prime example (Hinds, Williams, and Joinson, 2020; Wagner, 2021). This incorporation of a market logic into Canadian politics has transformed the way election campaigns are conceived and conducted (Marland et al., 2012). Political parties are now able to identify specific voters with a high potential for electoral return – a practice that, according to many, poses a serious threat to democratic vitality (Cwalina, Falkowski, and Newman, 2011; Johansen, 2012; Savigny, 2008; Scammell, 1999). Indeed, applied to the political world, the marketing logic induces a paradigm shift, where ideology loses its central importance. Nourished by research on the electoral market, political actors offer a "product" (Butler and Collins, 1999) that citizens "consume" by expressing their support, most often in the form of votes (see also Delacourt, 2013). In this context, many fear that politics is presented as a product rather than as a shared deliberation on social and political issues (e.g., see Maarek, 2007). Other researchers also point out the overemphasis on certain segments of the electorate at the expense of the common good or the concentration of political decisions in the hands of a few consultants (Lathrop, 2003; Lilleker, 2005; Nimmo, 1999). In this context, data-driven political marketing strategies by political parties are not without any risks for contemporary democracies.

These transformations also take place within a broader context of political disaffection and cynicism, as we observe a declining turnout over recent years combined with a decrease in people's confidence in political institutions – and Quebec and Canada are no exception (Dubois and Gélineau, 2021; Howe, Johnston, and Blais, 2005; Manoliu and Sullivan, 2016; Nadeau, 2002). Recent evidence suggests Canada has experienced a surge in partisan sorting since the early 1990s, resulting in consolidated partisan identifications (Kevins and Soroka, 2017) and a growing animosity between partisan groups (Cochrane, 2015; Johnston, 2019). It is probably reasonable to ask ourselves if the fact that politics is increasingly driven by data – and a logic of electoral gains – is not at least partly responsible for these transformations. After all, some researchers talk of a "marketing malaise" to explain some democratic challenges that electoral democracies, including ours, actually face (see Marland, Giasson, and Lees-Marshment, 2012). That being said, everything suggests that these practices will continue, and rigorous data governance on the part of social scientists is even more crucial. In this context, researchers also have a crucial role to play in making citizens aware of the wealth and value of these data – data that, as we know, are more likely to be used for electoral or marketing rather than for research purposes.

Already twenty years ago, Dalton (2000) was suggesting that the expanding collection of empirical data on public opinion was one of the major accomplishments in comparative politics (913). He was probably right, as long as political scientists know how to make sense of these data in a thoughtful and informed way. Political scientists – and public opinion analysts more specifically – need to find a balance between respecting citizens' right to privacy and collecting data that are now necessary to adequately study contemporary political communication and public opinion dynamics. Finally, in spite of the craze for new digital data, it is crucial that social observation remains guided by conscious theorizing. This is how the potential of digital data will be fully exploited.

Notes

[1] The "datafication" of the social world means that all its elements can be a continual source of data – "to datafy a phenomenon is to put it in quantified form so that it can be tabulated and analyzed" (Mayer-Schönberger and Cukier, 2013, p. 78).

[2] However, the digital divide creates new challenges in terms of data representativeness. Indeed, some groups are online less often, if at all, and are therefore more likely to be excluded from research. We can think of elderly people as an example or subgroups living in wireless death zones.

[3] On that note, these digital traces are increasingly used by companies and are the basis of "surveillance capitalism" (Zuboff, 2019).

[4] However, contrary to many sociodemographic characteristics, lifestyle is both culture-specific and not necessarily stable over time. Lifestyle is a fuzzy concept, and it is influenced by myriad intrinsic (e.g., personality traits) but also extrinsic (e.g., marketing) factors. Researchers dealing with this concept must keep in mind that their object of study per se requires a constant renewal and that the boundaries of lifestyle might need to be drawn and redrawn again.

[5] With Simon Coulombe, professor in the Department of Industrial Relations at Laval University.

[6] See www.datagotchi.com.

[7] Datagotchi needs to be distinguished from voting advice applications (VVAs), although both applications present some similarities – such as their educational

vocation and their interactive aspect. VVAs – we can think of Vote Compass, the most well-known and established VVA in Canada – are designed to *help* citizens manage the political landscapes by, for instance, showing them which party candidate best represents their preferences. Their goal is not to *predict* vote choice, especially since we know that most voters do not vote based on issue preferences (Lewis-Beck et al., 2009; Gidengil et al., 2012). Datagotchi, on the other hand, tries to guess people's vote choice based on nonpolitical questions. The objective is not to help people vote, but rather to make them think about the value of their personal data and the fact that they may be more predictable than they think.

8 The main objectives of the platform were communicated and explained to the users through the media. They also appear, in a few words, on the front page of the application. Although, for now, Datagotchi does not fit the criteria of big data (neither by their velocity, their volume, or their variety), the tool presents certain challenges (technical, theoretical, and ethical) inherent to the digital age.

9 MusicBrainz and OMDb more specifically. They are two open-source databases that contain metadata about musical artists and movies. Through an application programming interface (API), users' answers are linked to these databases. For instance, if a user indicates that their favourite musician/music group is The Beatles, the API will use MusicBrainz to fetch additional information, such as the genre, years of activity, names of albums, and more.

10 In the United States particularly, the relative lack of legislative constraints regarding campaign financing facilitates the advanced development of marketing tools and access to a wide range of data from private companies for political use.

References

Aral, S., and Walker, D. (2012). Identifying influential and susceptible members of social networks. *Science*, *337*(6092), 337–41.

Askitas, N., and Zimmermann, K.F. (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly*, *55*(2), 107–20.

Barberá, P., Wang, N., Bonneau, R., Jost, J.T., Nagler, J., Tucker, J., and González-Bailón, S. (2015). The critical periphery in the growth of social protests. *PloS one*, *10*(11), e0143611.

Beauchamp, N. (2017). Predicting and interpolating state-level polls using Twitter textual data. *American Journal of Political Science*, *61*(2), 490–503.

Bennett, W.L. (1998). The uncivic culture: Communication, identity, and the rise of lifestyle politics. *PS: Political Science and Politics*, *31*(4), 741–61.

Berger, P.L., Berger, B., and Kellner, H. (1973). *The homeless mind: Modernization and consciousness*. Anchor Books.

Bishop, B. (2008). *The big sort: Why the clustering of like-minded America is tearing us apart*. Houghton Mifflin Harcourt.

Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science*, *58*(2), 367–86.

Boyd, D., and Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication and society*, *15*(5), 662–79.

Butler, P., and Collins, N. (1999). A conceptual framework for political marketing. In B.I. Newman (Ed.), *Handbook of political marketing* (55–72). Sage.

Chaney, D.C. (2002). *Lifestyles*. Routledge.

Chetty, R., Hendren, N., and Katz, L.F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, *106*(4), 855–902.

Choi, H. and Varian, H. (2009a). *Predicting initial claims for unemployment benefits*. http://research.google.com/archive/papers/initialclaimsUS.pdf.

Choi, H. and Varian, H. (2009b). *Predicting the present with Google Trends*. http://google.com/googleblogs/pdfs/google_predicting_the_present.pdf.

Cinelli, M., Morales, G.D.F., Galeazzi, A., Quattrociocchi, W., and Starnini, M. (2020). Echo chambers on social media: A comparative analysis. arXiv:2004.09603. https://arxiv.org/abs/2004.09603.

Cochrane, C. (2015). *Left and right: The small world of political ideas.* McGill-Queen's University Press.

Coppersmith, G., Dredze, M., and Harman, C. (2014). Quantifying mental health signals in Twitter. In P. Resnik, R. Resnik, and M. Mitchell (Eds.), *Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality*, 51–60.

Couper, M. (2017). New Developments in Survey Data Collection. *Annual Review of Sociology 43*, 121–45.

Cwalina, W., Falkowski, A., and Newman, B.I. (2011). *Political marketing: Theoretical and strategic foundations*. ME Sharpe.

Dalton, R.J. (1984). Cognitive mobilization and partisan dealignment in advanced industrial democracies." *The Journal of Politics 46*(1): 264–84.

Dalton, R.J. (2000). Citizen attitudes and political behavior. *Comparative political studies*, *33*(6–7), 912–40.

Dalton, R.J., Wattenberg, M.P., and Finifter, A.W. (1993). *Political science: The state of the discipline II*. American Political Science Association.

De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., and Kumar, M. (2016). Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, 2098–110.

Delacourt, S. (2013). *Shopping for votes: How politicians choose us and we choose them*. D and M Publishers.

DellaPosta, D., Shi, Y., and Macy, M. (2015). Why do liberals drink lattes? *American Journal of Sociology*, *120*(5), 1473–511.

de Moor, J. (2017). Lifestyle politics and the concept of political participation. *Acta Politica*, *52*, 179–97. https://medialibrary.uantwerpen.be/oldcontent/container2608/files/ap201527a.pdf.

de Moor, J., and Verhaegen, S. (2020). Gateway or getaway? Testing the link between lifestyle politics and other modes of political participation. *European Political Science Review*, *12*(1), 91–111.

Dodds, P.S., Harris, K.D., Kloumann, I.M., Bliss, C.A., and Danforth, C.M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PloS one*, *6*(12), e26752.

Dubois, P.R., and Gélineau, F. (2021). Les motifs de la participation électorale aux élections municipales québécoises: le cas de 2017. [Research Report] Chaire de recherche sur la démocratie et les institutions parlementaires, Université Laval.

Dufresne, Y., and Marland, A. (2012). The Canadian political market and the rules of the game. In A. Marland, T. Giasson, and J. Lees-Marshment (Eds.), *Political marketing in Canada* (22–38). University of British Columbia Press.

Duncombe, C. (2019). The politics of Twitter: Emotions and the power of social media. *International Political Sociology*, *13*(4), 409–29.

Fischer, C.S., and Mattson, G. (2009). Is America fragmenting? *Annual Review of Sociology*, *35*, 435–55.

Flanagan, T. (2009). *Harper's team: Behind the scenes in the Conservative rise to power*. McGill-Queen's University Press.

Florida, R. (2008). *Who's your city? How the creative economy is making where to live the most important decision of your life*. New York: Basic Books.

Gidengil, E., Blais, A., Everitt, J., Fournier, P., and Nevitte, N. (2012). *Dominance and decline: Making sense of recent Canadian elections*. University of Toronto Press.

Franklin, C.H. (1992). Measurement and the dynamics of party identification. *Political Behavior*, 297–309.

Gendreau, P. (2023). GAFAM, Le monstre à cinq têtes. Éditions Radar.

Giddens, A. (1991). *Modernity and self-identity: Self and society in the late modern age*. Stanford University Press.

Goel, S., Hofman, J.M., Lahaie, S., Pennock, D.M., and Watts, D.J. (2010). Predicting consumer behavior with Web search. *Proceedings of the National Academy of Sciences*, *107*(41), 17486–90.

Goldberg, A. (2011). Mapping shared understandings using relational class analysis: The case of the cultural omnivore reexamined. *American Journal of Sociology*, *116*(5), 1397–436.

Hetherington, M., and Weiler, J. (2018). *Prius or pickup? How the answers to four simple questions explain America's great divide*. Houghton Mifflin.

Hinds, J., Williams, E.J., and Joinson, A.N. (2020). "It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human-Computer Studies*, *143*, 102498.

Howe, P., Johnston, R. and A. Blais (Eds.). (2005). *Strengthening Canadian democracy*. Institute for Research on Public Policy.

Huber, J.D., and Suryanarayan, P. (2016). Ethnic inequality and the ethnification of political parties: Evidence from India. *World Politics*, *68*(1), 149–88.

Isitor, E., and Stanier, C. (2016). Defining big data. In *Proceedings of the International conference on big data and advanced wireless technologies*, 1–6.

Jackson, S.J., and Foucault Welles, B. (2016). #Ferguson is everywhere: Initiators in emerging counterpublic networks. *Information, Communication and Society*, *19*(3), 397–418.

Johansen, H.P. (2012). *Relational political marketing in party-centred democracies: Because we deserve it*. Ashgate Publishing.

Johnston, R. (2019, June 4–6). Affective polarization in the Canadian party system, 1988–2015. Annual Meeting of the Canadian Political Science Association. Vancouver, BC, Canada.

Judge, E.F., and Pal, M. (2021). Voter privacy and big-data elections. *Osgoode Hall LJ*, *58*, 1.

Kevins, A., and Soroka, S.N. (2017). Growing apart? Partisan sorting in Canada, 1992–2015. *Canadian Journal of Political Science / Revue canadienne de science politique*, *51*(1), 103–33.

Lamont, M., Schmalzbauer, J., Waller, M., and Weber, D. (1996). Cultural and moral boundaries in the United States: Structural position, geographic location, and lifestyle explanations. *Poetics*, *24*(1), 31–56.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, *6*(70), 1.

Lathrop, D.A. (2003). *The campaign continues: How political consultants and campaign tactics affect public policy.* Praeger.

Lazarsfeld, P.F., Berelson, B., and Gaudet, H. (1944). *The people's choice.* Duell, Sloan and Pearce.

Lazer, D., Kennedy, R., King, G., and Vespignani, A. (2014). The parable of Google Flu: Traps in big data analysis. *Science*, *343*(6176), 1203–205.

Lazer, D., and Radford, J. (2017). Data ex machina: Introduction to big data. *Annual Review of Sociology*, *43*, 19–39.

Leblanc, D. (2021). Ottawa abandonner l'utilisation de la Libéraliste pour la sélection des juges. *Radio-Canada*. https://ici.radio-canada.ca/nouvelle/1800125/liberaliste-selection-juge-gouvernement-federal.

Lewis-Beck, M.S., Norpoth, H., Jacoby, W.G., and Weisberg, H.F. (2009). *The American voter revisited*. University of Michigan Press.

Lilleker, D.G. (2005). The impact of political marketing on internal party democracy. *Parliamentary Affairs*, *58*(3), 570–84.

Lipset, S.M., and Rokkan, S. (1967). *Party systems and voter alignments: Cross-national perspectives*. Free Press.

Litle, P., and Zuckerman, M. (1986). Sensation seeking and music preferences. *Personality and individual differences*, 7(4), 575–78.

Maarek, P.J. (2007). *Communication et marketing de l'homme politique*. Litec.

MacDonald, R., Hargreaves, D.J. and Miell, D. (2017). *Handbook of musical identities*. Oxford University Press.

Manoliu, I.A., and Sullivan, K.V. (2016). Youth participation and cynicism during the Canadian federal election of 2015. [Research report], Élections Canada.

Marland, A., Giasson, T. and Lees-Marshment, J. (2012). *Political marketing in Canada.* University of British Columbia Press.

Mayer-Schönberger, V., and Cukier, K. (2013). *Big data: A revolution that will transform how we live, work and think.* John Murray.

Michel, J.B., Shen, Y.K., Aiden, A.P., Veres, A., Gray, M.K., Google Books Team, Pickett, J. P, Hoiberg D., Clancy D., Norvig P., Orwant J., Pinker S., Nowak M.A. and Aiden, E.L. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, *331*(6014), 176–82.

Micheletti, M., and Stolle, D. (Eds.). (2004). *Politics, products, and markets: Exploring political consumerism past and present*. Transaction Publishers.

Mutz, D.C., and Rao, J.S. (2018). The real reason liberals drink lattes. *PS: Political Science and Politics*, *51*(4), 762–67.

Nadeau, R. (2002). Satisfaction with democracy: The Canadian paradox. In N. Nevitte (Ed.), *Value change and governance in Canada* (37–70). University of Toronto Press.

Nickerson, D.W., and Rogers, T. (2014). Political campaigns and big data. *Journal of Economic Perspectives*, *28*(2), 51–74.

Nimmo, D. (1999). The permanent campaign: Marketing as a governing tool. In B.I. Newman (Ed.), *Handbook of political marketing* (73–86). Sage.

Norris, P., and Mueller, C.M. (1988). *The politics of the gender gap: The social construction of political influence*. Sage.

North, A.C., and Hargreaves, D.J. (2007). Lifestyle correlates of musical preference: Relationships, living arrangements, beliefs, and crime. *Psychology of music*, *35*(1), 58–87.

Onnela, J.P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J., and Barabási, A.L. (2007). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, *104*(18), 7332–336. https://www.pnas.org/doi/full/10.1073/pnas.0610245104

Page, S.A., and Nyeboer, J. (2017). Improving the process of research ethics review. *Research Integrity and Peer Review*, *2*(1), 1–7.

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., and Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, 1.

Plasser, F., and Plasser, G. (2002). *Global political campaigning: A worldwide analysis of campaign professionals and their practices*. Praeger.

Polgreen, P.M., Chen, Y., Pennock, D.M., Nelson, F.D., and Weinstein, R.A. (2008). Using internet searches for influenza surveillance. *Clinical Infectious Diseases*, *47*(11), 1443–448.

Rentfrow, P.J., and Gosling, S.D. (2007). The content and validity of music-genre stereotypes among college students. *Psychology of Music*, *35*(2), 306–26.

Rentfrow, P.J., McDonald, J.A., and Oldmeadow, J.A. (2009). You are what you listen to: Young people's stereotypes about music fans. *Group Processes and Intergroup Relations*, *12*(3), 329–44.

Rose, R. (1974). *The problem of party government*. Free Press.

Rose, R. and McAllister, I. (1986). *Voters begin to choose: From closed-class to open elections in Britain*. Sage.

Saunders, P. (1981). *Social theory and the urban question*. Routledge.

Saunders, P. (1990). *A Nation of home owners*. Unwin Hyman.

Savigny, H. (2008). *The problem of political marketing*. Continuum.

Scammell, M. (1999). Political marketing: Lessons for political science. *Political studies*, *47*(4), 718–39.

Shor, E., Van De Rijt, A., Miltsov, A., Kulkarni, V., and Skiena, S. (2015). A paper ceiling: Explaining the persistent underrepresentation of women in printed news. *American Sociological Review*, *80*(5), 960–84.

State, B., Park, P., Weber, I., and Macy, M. (2015). The mesh of civilizations in the global network of digital communication. *PloS one*, *10*(5), e0122543.

Stolle, D., Hooghe, M., and Micheletti, M. (2005). Politics in the supermarket: Political consumerism as a form of political participation. *International political science review*, *26*(3), 245–69.

Toole, J.L., Lin, Y.R., Muehlegger, E., Shoag, D., González, M.C., and Lazer, D. (2015). Tracking employment shocks using mobile phone data. *Journal of The Royal Society Interface*, *12*(107), 20150185.

van de Rijt, A., Shor, E., Ward, C., and Skiena, S. (2013). Only 15 minutes? The social stratification of fame in printed media. *American Sociological Review*, *78*(2), 266–89.

Van Mol, Christof. (2017). Improving web survey efficiency: The impact of an extra reminder and reminder content on web survey response. *International Journal of Social Research Methodology*, *20*(4), 317–27.

Wagner, P. (2021). Data privacy-the ethical, sociological, and philosophical effects of Cambridge Analytica. SSRN. http://dx.doi.org/10.2139/ssrn.3782821.

Wang, W., Rothschild, D., Goel, S., and Gelman, A. (2015). Forecasting elections with non-representative polls. *International Journal of Forecasting*, *31*(3), 980–91.

Weiss, M.J. (1988). *The clustering of America*. Harper and Row.

Wilson, H. (2018). The "I" in AI: Emotional intelligence and identity in Ex Machina. *Film Matters*, *9*(1).

Zeitzoff, T. (2017). How social media is changing conflict. *Journal of Conflict Resolution*, *61*(9), 1970–991.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.

Zweigenhaft, R.L. (2008). A do re mi encore: A closer look at the personality correlates of music preferences. *Journal of individual differences*, *29*(1), 45–55.

6

Party Members, Canvassing, and Microtargeting

Ethnography of a Data-Driven Campaign in Turin, Italy

Cecilia Biancalana

Introduction

Data has become an important part of our understanding of electoral campaigns (Dommett, 2019); more and more, political parties use data to deliver tailored messages to their audiences. Although a gap has been detected between the rhetoric and practices of data-driven campaigns (Baldwin-Philippi, 2017) – meaning that their reality is often less sophisticated than journalistic coverage may make us think – electoral campaigns based on the collection and strategic use of various types of data are here to stay. And this trend is likely to evolve and amplify with the use of AI by political parties.

While the promises of this type of campaigning include reaching social groups difficult to contact and giving citizens information about the campaign in a more efficient and effective manner (Zuiderveen Borgesius et al., 2018), the risks for democracy of these kinds of practices lie in privacy intrusion (Bennett and Lyon, 2019; see Bennett, this volume), manipulation of the vote, and commercialization and exclusion of those not able to participate in the digital world (Gorton, 2016).

Beyond the major issues linked to the democratic implications of these practices (see Blais, this volume), an equally important but underdeveloped question is *how, concretely,* parties obtain data to orient their campaign strategies and convey a tailored message to the electorate. It has been shown (Dommett, 2019) that it may be through emails, social media, and public records but also in-person canvassing (Nielsen, 2012).

Indeed, although to some extent paradoxically, it is precisely in an era characterized by the massive use of digital media that on-the-ground and grassroots campaigns regained their importance; canvassing now increasingly involves the use of data to help party members or volunteers to reach their targets and also tends to produce data on electors. Moreover, experts are not always behind the collection of data for microtargeting; parties often rely on the activism of local volunteers and not on expert data professionals (Dommett, 2019).

As data-driven campaigning is context-dependent, and less efficient and professional than one may think, and as its strategies and practices may diverge, it is important to study its concrete unfolding. In particular, as data are the raw material upon which microtargeting campaigns are based, looking at its collection could help us to better understand the risks (or potential benefits) of these practices for democracy.

Against this backdrop and to contribute to the literature on the concrete practices of data-driven campaigns, my aim in this chapter is to analyze from the inside, with ethnographic methods, an electoral campaign carried out in the city of Turin, Italy, in 2016, named Noi Siamo Torino (NST). NST was an electoral campaign inspired by the US model and based on volunteers' field mobilization and microtargeting. It was the first campaign of this kind carried out in a large Italian city. The campaign was aimed at the re-election of Mayor Piero Fassino, a member of the Partito Democratico (PD), supported by a centre-left coalition.

As Fassino lost the election in the second round to the Five Star Movement's (M5S) candidate, Chiara Appendino, the case of NST will not be analyzed here from the point of view of electoral effectiveness (see Cepernich, 2017, 105–6). On the contrary, through the analysis of the NST campaign, based on direct observation and semi-structured interviews, the goal of this contribution will be to show "what is behind" data collection in an electoral campaign based on field mobilization and microtargeting. This is of the utmost importance because ethnographic methods can allow us to enter the "secret garden" of politics and to look concretely and comprehend practices, such as microtargeting and (digital) campaigning, often studied with quantitative approaches, when not kept secret by parties themselves.

This case study will allow us to examine the importation of data-driven campaigns in the European context, in which political parties have a different role with respect to the US. NST was indeed an organization, external and separate from the candidates' party, that had two main objectives: recruit "volunteers" with the task of persuading citizens to vote for the candidate, Fassino, using their experience of lay citizens; and collect, through volunteers' activities, data for microtargeting. The two goals represent a novelty for an Italian (or European) party: on the one hand, the creation, through the activity of volunteers, of a direct relationship between citizens and candidates that was supposed to encourage an increased turnout; on the other hand, the interactions also had the aim of collecting data for microtargeting.

Results show that although the goal of the organization was to mobilize lay citizens, in the vast majority of cases, PD party members took part in the campaign as "volunteers." Although the organization gave precise instructions to volunteers regarding how to interact with citizens and to collect data for microtargeting, volunteers often diverged from party instructions, also with respect to data collection. In sum, the examination of the campaign gives a different account with respect to the common understanding of data-driven campaigns and helps us to better understand the role of party members and their resistances to new campaign practices.

Where Do Data-Driven Campaigns Come From?

Studies on the evolution of political communication in general (Blumler and Kavanagh, 1999), and of electoral campaigns in particular (Norris, 2000), highlight the succession of different phases in the evolution of electoral campaigns. Indeed, over the last decades, some social, political, and technological transformations changed the way of conducting and communicating politics. This is crucial for understanding the emergence of campaigns based on microtargeting.

It is possible to identify three (for some, four, see Blumler, 2016) main phases. The first, which goes from the postwar period to the 1950s, is marked by the supremacy and hegemony of parties, especially in Europe. In a context in which political alternatives seem "frozen" (Lipset and Rokkan, 1967) and in which citizens' identification with political parties is high, parties are the main actors of political communication and electoral campaigns, in part because of the control they can exercise over the media. In this phase, electoral campaigns are managed independently by the party through *direct forms of interaction between candidates and voters* (Manin, 1995) and propaganda broadcast through partisan media.

The second phase, which can be said to have begun in the 1960s, coincides with two trends: the diffusion of television and the loosening of party loyalty. These are two crucial changes that radically transformed political communication and electoral campaigns. The result is the spread of mediatization processes (Mazzoleni and Schulz, 1999): an increasing influence of the so-called media logic (Altheide and Snow, 1979) in all fields, including politics. Political institutions are increasingly dependent and shaped by the media, especially by television, rather than vice versa.

A greater independence and commercialization of the media imply a different way of dealing with politics, one that is more oriented towards forms of spectacularization, popularization, and fusion with entertainment (van Zoonen, 2005). *Electoral campaigns are increasingly played out on television*, and therefore their organization is increasingly entrusted to professionals who do not belong to the party. Since a specific expertise is needed to manage this kind of campaign, and because of declining numbers, the party relies less and less on members for mobilization and propaganda activities (Dalton and Wattemberg, 2002; van Biezen, Mair and Poguntke, 2012). Therefore, with the passage from the first to the second phase, *the personal and direct contact between citizens and candidates loosened* (Lefebvre, 2016) and the distance between citizens and politicians grew (Manin, 1995).

An opposite tendency seems to have emerged in the current phase. The massive diffusion of digital media, especially those relating to the so-called web 2.0, represent a deep paradigm shift compared to the past, defined by Manuel Castells (2009) as "mass self-communication." On the one hand, with the internet, citizens have a greater responsibility in the selection of information; on the other, they can also become producers and distributors of information, bypassing journalistic mediation.

Information therefore becomes more individualized and personalized and potentially more horizontal and bidirectional. At the same time, political identities and parties' organizational structures further transformed into postbureaucratic forms. Light and multispeed forms of participation and affiliation emerged (Scarrow, 2014). Furthermore, the massive diffusion of digital media and the decline of the number of party members and the workforce they represented fostered the systematic and strategic use of data by parties during electoral campaigns.

In this context, over the past fifteen years, starting from the United States, electoral campaigns that enhanced *direct contact with the voters*, through telephone and personal contact, appeared. It is what Nielsen (2012) defines as "personalized political communication": organized practices that use people as media for political communication. In reality, the United States has a vast tradition of canvassing that never disappeared. But, starting from the 1990s, and especially around 2000, we witness a growth in this practice, which we can attribute to the diffusion of the research by two political scientists from Yale, Alan S. Gerber and Donald P. Green. Gerber and Green (2000, 2004), through experimental research, demonstrated the electoral effectiveness of canvassing,[1] especially regarding voters that would not have otherwise voted (hence the term "mobilization campaign," i.e., that does not aim to convince voters already inclined to vote for the other candidate but to mobilize as much as possible the voters belonging to their own field), because of the social pressure that the campaign would create.

Canvassing is indeed one of the ways through which parties can collect data on voters, and canvassing can be also guided by data. According to Bodó, Helberger, and de Vreese (2017), political microtargeting "refers to the use of different communications (mail, phone, canvassing, direct mail, and social media advertising, etc.) to communicate and build a relationship with prospective voters. At the core of the concept is the use of data and analytics to craft and convey a tailored message to a subgroup or individual members of the electorate."

The systematic and strategic use of data, the construction of databases, and the analysis and segmentation of data to profile voters and contact them are features that characterize modern data-driven campaigns. These features had an impact also on canvassing.

Although there are some interesting previous experiences (for example, the Republican election campaign for the US presidential elections in 2004), it is Obama's 2008 campaign that made microtargeting globally famous and became the reference point for all future experiences of this kind. The 2008 US presidential election was indeed the first major election that perfected microtargeting models, using information about voters to direct volunteers to scripted conversations at the door or over the phone.

As noted by Lefebvre (2016), Obama's victory helped to legitimize the use of these new practices and to "export" them, even despite the different characteristics of the European and American social and political contexts and consequently the different structures of their electoral campaigns. In this respect, an interesting importation of this kind of experience in the European context is the French one, which took place during the 2012 presidential elections.[2] Three young French scholars affiliated to US universities, after having participated in the Obama campaign in 2008, proposed a similar campaign to the French Socialist Party. The three proponents, called *Les Bostoniens*, used Socialist Party's members to organize a massive national canvassing campaign (Lefebvre, 2016; Liégey, Muller, and Pons, 2013; Talpin and Belkacem, 2014), reinventing and innovating in the direction of greater managerialization, not without resistances, the political communication practices of a Socialist Party affected by a membership decline.

In conclusion, the fundamental characteristic of data-driven mobilization campaigns is the *union of the possibilities offered by new technologies with the search for personal and direct contact with voters* or, rather, the attempt to organize in a scientific and managerialized way, also through data, direct contact with voters, as was characteristic of the first phase of political communication.

From the US to Turin

The data-driven mobilization campaign that took place in Turin in spring 2016 was the first of its kind organized in a large Italian city. The organization of NST was autonomous and detached with respect to the PD but specifically aimed at the re-election of the incumbent mayor Fassino. NST was only a part of Fassino's electoral campaign, and certainly a minor part compared to the economic investment by the candidate.

Fassino relied largely on the classic electoral campaign tools such as billboards and on the network of the PD's sections that organized stands and other events during the electoral campaign. NST was therefore an autonomous organization, integrated into Fassino's electoral campaign.

How was a US electoral campaign "imported" into the Turin context? The creators and coordinators of the project were Cristopher Cepernich, a professor of sociology at the University of Turin and an expert in political communication, and Flavio Arzarello, an expert in political communication and a party executive of the PD. Cepernich and Arzarello proposed the project to Fassino who, in a context in which the party was suffering from a deep membership crisis (Natale and Fasano, 2017), accepted, hoping to obtain an additional mobilization.

The two creators of the project showed a great fascination for American election campaigns. Cepernich was present as an observer in Lecco, a city in which a similar campaign for the re-election of the mayor was managed by Mike Moffo, media strategist in Obama's 2012 electoral campaign. Cepernich and Arzarello met at a conference at which Moffo was a speaker, organized by Arzarello in 2014. Even in this case, as in the French one, there seems to have been a "contagion"[3] from the United States: the reference to the US experience and to Obama is, in fact, explicit.[4] It is also important to consider that academia played an important role in this case, as it also did in the US and in France: Cepernich, in fact, played the dual role of scholar and political consultant.

Members of the NST staff (about fifteen to twenty people) were also university students. They were students of public and political communication at the University of Turin interested in experiencing an electoral campaign on the field. They were very young, all under the age of thirty. Their relationship with the electoral campaign was of a formative and professional nature. The campaign is considered a valuable experience in the field of political consultancy, to be "added to one's CV," regardless of the candidate's political colour ("if they had called me in Milan to campaign for Berlusconi, I would have gone" – one of the staffers told me).

The staff's task, under the direction of Cepernich and Arzarello, was to organize and coordinate the mobilization of volunteers. Indeed, in the minds of the creators, NST was to stand on two legs. On the one hand, the goal was the mobilization of "volunteers," with the task of (according to the presentation slides of the project) "reactivating the nodes of relationship in the territory," "reactivating participation and increasing turnout," and "promoting the desired voting behavior" through direct and interpersonal communication with voters. On the other hand, volunteers would have to collect data on voters "for the communication and the strategy of the campaign."

Before I analyze these two goals in detail, the use of the term "volunteers" deserves further attention. In Italy, unlike in other countries such as the United States, the term "volunteer" refers to the semantic field of social volunteering. By defining the volunteers as such, it seems that the goal is to untie the volunteers for Fassino from the political dimension and to link them to civic activism.

The relationship between volunteering and politics in Italy has been marked by different phases (Biorcio, Caruso, and Vitale, 2016). Already in the 1960s, the research by Almond and Verba (1963) showed that, in Italy, the figures for membership in associations were lower than in other European countries and that civic associations had strong ties with mass parties. Only between the end of the 1960s and the 1980s, with the decrease in party identification and especially after the collapse of the Italian party system in the early 1990s (Grilli di Cortona, 2007), civic associations started to gain autonomy from parties.

From Toqueville onwards, participation in associations has always been considered a "school of democracy," capable of socializing citizens to democratic practices and increasing participation and citizens' trust in institutions. However, although even political activism can be considered a voluntary activity, the two types of participation differ substantially. While social participation is situated within civil society, and its aim is to protect rights and common goods and to support disadvantaged individuals outside the actions of parties (Moro, 2013), party membership and activism belong to the political field, and we can say that their aim is to influence the "authoritative allocation of values in a society" (Easton, 1965) through the election of candidates to public offices. While active citizenship tends to be self-organized, participation in parties takes place within hierarchical structures organized from above (Michels, 1911).

We can say then that NST tried to import some elements of social volunteering, or more generally of active citizenship, into the political field. It is therefore possible to define this campaign as a depoliticized one, meaning that we can see a shift of campaign practices from the political to the personal-social dimension (Flinders and Wood, 2014; Wood and Flinders, 2014).

Coherently with this frame, in the project's presentations, NST volunteers are depicted as separated and disconnected from the PD. The goal of volunteers should be to interact with and listen to citizens, demonstrating – through their presence on the ground – the closeness of the candidate to the citizens. Nevertheless, although one of the aims of the campaign was to mobilize lay citizens, especially those not necessarily active in politics and the party, the promoters also sought to structure and organize their mobilization with the collaboration of the local sections of the PD (*circoli*).

Volunteers could in fact register through a dedicated website and were coordinated and managed by NST, but the local sections of the PD were also supposed to play a part in the campaign. Thanks to their territorial rooting, the local sections were meant to represent a sort of "logistic base" for NST volunteers and contribute to the campaign with activists who knew their neighbourhood and could therefore help volunteers. During the presentations of the project in the local sections, organizers told party members that NST would do "what parties have always done" but with a strategy and in a systematic way.

However, in reality, the organizers' idea was to structure in different ways the campaigns of the party and of the volunteers. At the meetings that took place in the early phase of the campaign in various local sections of the party, the organizers presented a clear "division of labour" between the two types of actors involved in the campaign: NST volunteers, unlike PD members, should establish a relationship with citizens based on their *personal and subjective* experience of the city and not on the basis of an electoral program or party identity.

But the NST campaign was not only based on the mobilization of volunteers and on their direct interaction with citizens. The other side of the campaign was the systematic collection by volunteers of voters' data, which were used to direct the electoral campaign. In fact, the conversation with the citizens was not intended as an end in itself. The aim was to fill a form that, in addition to a voter's personal data, contained two pieces of information that would serve to profile the voter and to send them targeted communication: their priorities for Turin and their voting orientation.

The collection of data by volunteers could take place in three different ways: first, through the volunteers' private relationships (e.g., relatives, acquaintances); second, through their presence in the neighbourhoods' public places; and finally, through canvassing. In the last two cases, the volunteers were equipped with a red bag and a red bib printed with the words "Piero Fassino candidate mayor of Turin" and instructed to conduct a "completely natural" conversation based on the following script:

> Good morning/good evening, my name is Chiara, I am a volunteer with Noi Siamo Torino for the electoral campaign of Piero Fassino. I would like to talk to you briefly about our city. Are you available?
>
> If so:
>
> In June there will be the elections for the mayor of Turin. Have you already decided who to vote for? (If you notice enthusiasm, ask if the person is willing to dedicate a part of his or her time to help as a volunteer.) In your opinion, what is the most urgent problem that the next mayor of Turin will have to solve? Do you agree to give us your contact details to establish regular contacts and receive detailed information about our activities and our program?
>
> If not (quickly understand why):
>
> If he or she is adamant about voting for another candidate, quickly close the conversation and move on. If he or she fears revealing his or her opinions, be reassuring and underline the VOLUNTARY aspect of your work.

What Is behind Microtargeting

"Hello, we are volunteers for Piero Fassino; can we ask you some questions about the city?" This was the typical beginning of an interaction between NST volunteers and a citizen of Turin. We have seen how, behind this simple question, there was a wider strategy and project, which can be linked with the transformations of electoral campaigns and party membership. In the end, organizers declared that they had collected 11,507 forms for microtargeting. However, an organization's strategy, which appears to be highly rationalized, is often different from the actual practices implemented by militants and volunteers (Belkacem and Talpin, 2014). How, thus, did the campaign unfold?

To answer this question, I took part in the activities of the group for four months, from March to June 2016, clearly declaring the purpose of my participation. During this period, I attended the weekly staff meetings and took part in the activities of the volunteers on the ground, mainly but not exclusively in a particular district that was semi-peripheral and mainly residential. I participated in all electoral campaign activities (rallies, flyer handouts, etc.) and meetings. I also took part in some public events of the electoral campaign the group was involved in. I had complete access to the field operations, to staff meetings, and to the WhatsApp chat of the group. I did not conduct interviews with the electoral campaign staff; my interactions with them took place mainly "on the ground" while they were doing their job. However, I did interview ten campaign volunteers with the purpose of investigating their social and political identity, their motivation for joining the party, and their perception of the electoral campaign.

As regards the campaign organization, to structure the mobilization in a capillary way throughout the city, eight small groups were created in correspondence to the eight administrative districts of Turin. Each group was managed by two staff members (called "captains"), who had the task of coordinating the operations of the volunteers. The operations mainly consisted of two types of "outings" (*uscite*): those in public places (e.g., markets) and door-to-door canvassing. The PD's local sections were supposed to serve as a logistic base.

A higher level of coordination was represented by the two campaign managers (Cepernich and Arzarello) and by another person in charge of the organization. This person had the task of contacting people who registered as volunteers through the website and sending them to the captains. All the outings were inserted in a Google calendar, and the staff was coordinated through a WhatsApp chat group. Once a week, the staff met at the headquarters of the electoral committee to discuss the progress of the campaign. It is at this central level that it was possible to monitor the progress of the microtargeting data collection. The data were processed using a software (Target 51) that allowed NST to profile voters and to send them targeted communication.

As regards volunteers, it is not possible to say exactly how many of them actually took part in the campaign, as the purpose of NST was precisely to mobilize a network of both formal and informal supporters. A leaflet distributed during a meeting between Fassino and the volunteers noted that volunteers can have different forms of commitment and involvement, ranging from taking part in the ground operations ("Come with us on the ground") to lighter forms of participation, such as speaking about NST to friends or following NST on social networks. The idea of involving personal networks is greatly emphasized ("introduce us to friends"). Therefore, in a broader sense, we can consider a volunteer anyone who solicited a vote for Fassino with phone calls. For the purposes of this research, however, we will only consider the most active volunteers, those who took part in the activities on the ground.

Thanks to the collaboration of the staff, I asked each captain to indicate the number of volunteers who had taken part in at least one outing on the ground. The number of volunteers turned out to be twenty-seven. I then asked the captains to indicate to me the volunteers who had taken part at least three times, and that number turned out to be twelve. Ten of the twelve active volunteers were interviewed.

To start, it is necessary to highlight that this number of volunteers is very low, in general and in relation to the population of Turin, which is around 900,000. This, against the broad visibility of the project, advertised through all the channels of Fassino's election campaign, indicates the low appeal of such an initiative in the Turinese context. Moreover, seven of the ten volunteers were members of the PD (the eldest were also members of the Partito Comunista Italiano – PCI, the mass party ancestor of the PD) and regularly participated in the activities of a PD local section. One volunteer was a member in the past, but, disagreeing with some of the party's choices, he has not renewed his membership; another is not a member but considered himself close to the party (see Table 6.1).

Table 6.1 **Summary of the profile of the interviewees**

No.	*Age*	*Sex*	*Profession*	*Member of the PD*	*Social Volunteering*
1	23	M	University student	Yes	Yes, animal rights
2	17	M	High school student	Yes	No, politics is volunteering
3	68	M	Retired teacher	Yes, from the PCI	Yes, local committees
4	46	M	Tailor	No, but close	Yes, immigrants' integration
5	27	F	University student	No, not identified with any party	Yes, social volunteering
6	70	M	Retired administrative executive	Yes, from the PDS	No, politics is volunteering
7	68	F	Retired janitor	Yes	Yes, social volunteering
8	23	M	University student	No, but he was Primary voter	No, but he would like to
9	55	M	Worker	Yes, from the PCI. Member of trade unions	Yes, social volunteering
10	78	M	Retired entrepreneur	Yes, from the PCI	No, politics is volunteering

We also asked the volunteers whether they were active in the field of social volunteering: six volunteers were, while three said that, for them, their "volunteering" activity was political ("I am already a volunteer in the party," Int. 2). In this sense, the position of volunteer no. 10, a long-time party member, is interesting because he suggests that party membership is understood by party members as a "voluntary" activity and that, for them, the political activity is perceived as more important than the social one ("We were *more* volunteers").

Therefore, despite the attempt to import some elements of volunteering and more generally of active citizenship into the political field, we note that the people interested in this project had, in the vast majority of cases, a very strong political and party background. We can thus compare volunteers' representations of the campaign with their other experiences of political activism. What emerges from the interviews, especially with regard to older militants, is in fact a contrast between the experience of NST, characterized by the presence of volunteers on the ground, and the experience of the PD in recent years, characterized by the closure of sections, the lack of members, and a growing detachment from the territory.

Figure 6.1 shows the evolution of the PD's membership figures compared to the participation in the party's primaries, both in the open phase and in that reserved to members. Indeed, the PD presents a peculiar organizational structure, which has been defined as "open" (Vassallo and Passarelli, 2016). There is a very nuanced distinction between party *supporters*, who have many rights, including the opportunity to vote for the party secretary (and not just for candidates) during party primaries, and *members*, who have few additional rights compared to supporters (for instance, a vote in the first phase of the party primaries[5]). The constant decrease in all the three figures testify to a crisis of mobilization of the party.

Against these difficulties, for older members, NST represents a sort of continuity with a "golden age" – with "what we did." For instance, PCI members used to deliver the party newspaper *L'Unità* door-to-door and had a strong rooting in the neighbourhoods. Generally, the presence of NST was perceived as the consequence of the absence of party members on the ground.

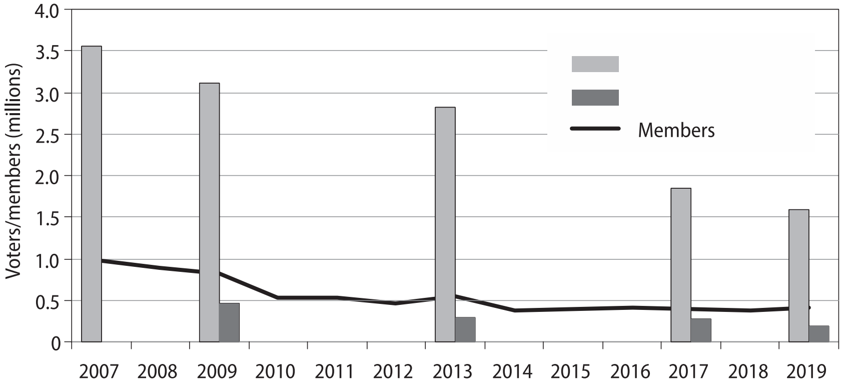

Figure 6.1 **PD's membership and voters in primaries (closed and open phases) |** Data courtesy of the Italian Democratic Party

> If this [NST] continues, it will become like what we did: every Sunday we sold *L'Unità* door-to-door. In the neighbourhood, they knew you because they always saw you. There was constant activity. That is what is missing now, in fact, because many local sections are always closed. (Int. 10)

If, on the one hand, members' perception is that NST "does what we have always done" but in a more structured way, for others, there is suspicion that volunteers were "paid": staff members were people unrelated to local sections, and it was difficult for members to understand why they were committed to Fassino and not to the PD. This testifies to the fundamental difference between Europe and the United States, where political volunteers, paid or not, are a common practice.

With regard to the outings, which took place in either public places or door-to-door, volunteers were supposed to talk with and listen to citizens, proposing their personal motivations for voting for Fassino and collecting data for microtargeting.

In the face of a highly rationalized strategy, it is interesting to analyze what happened, concretely, during the outings, how the volunteers interpreted the role they had been assigned, and how data was collected. A starting point is the general climate of negativity and opposition to politics among citizens. Faced with this climate of opinion, the strategy of volunteers was to shift the focus from the figure of Fassino (who, as incumbent mayor, is perceived as part of the establishment) to issues related to the city. "Do not immediately say that you are here for Fassino," was the advice given.

In short, what appeared to take place was an attempt to depoliticize the local campaign. "It's like selling a product, but here you sell a political idea," says a volunteer. The consequence of this approach, however – also given the absence of booths and gazebos typical of a party electoral campaign in Italy and the presence of bibs without party symbols – is that the volunteers were mistaken for ActionAid or Greenpeace volunteers, who ask for donations on the street and tend to be ignored by passersby.

Second, there is a discrepancy between the declared objectives of the project and their concrete realization. On the one hand, for staff members, younger and more aware of the mechanism behind the campaign, the goal seemed to be essentially to collect data rather than to interact with citizens. If, according to the coordinators, the form is the tool with which to create a relationship and not the very goal of the interaction, for some staff members, the goal seemed to be simply to fill out the form. On the other hand, volunteers, as we will see below, often struggled to understand the sense of the data collection. This clearly emerges in an exchange between a volunteer and a staff member:

> The volunteer (a man in his mid-sixties) asks the staff member what the goal is, whether to "collect data" or to "sow." The staff member answers without hesitation that the goal is to collect data. The volunteer tries to understand what these data are for; he is a little doubtful but interested. The staff member is interested in the data, and it is clear his goal is to collect as much data as possible, and in fact he runs swiftly from one interaction to another. Instead, the volunteer starts talking to people calmly. People reject him less because he is an older man, and he seems nice and quiet. He does what volunteers should do on paper: he asks people to talk about the city and listens to them. In fact, I notice that on some occasions this works; some people who initially do not want to talk eventually give their data. However, it does not always work; some people do not want to talk. "It's like

> selling a product, but here you sell a political idea," he tells me. The volunteer seems especially pleased to see a young man (the staff member) who engages in politics. But, as soon as the outing ends, the staff member tries to get rid of the red bag because he is going to the grocery store and does not want to go there with a bag with the name of Fassino on it. In the end, he turns the bag inside out, so that it is impossible to see the writing. (Fieldwork note, April 5, 2016)

The element of rejection and mistrust was also present during the canvassing. On these occasions, the volunteers were supposed to go from house to house to remind voters about the elections and to solicit a vote for Fassino (the so-called GOTV – Get Out The Vote). Especially when it was done in the absence of people belonging to the neighbourhood, or party members or candidates, there was a strong reluctance to open doors to strangers. In the absence of the party's previous presence in the neighbourhoods, the experiment seemed destined to fail.

> I'm going to canvass with two staff members. We have to walk the entire street, contacting all the people who vote in an electoral section. We have, in fact, a list of all voters residing in that street who vote in a certain section (the walk-list). One of the staff members tells me that he contacted the members of the Partito Democratico who live in that road to go in their building with them, but no one is available. We are on a street in a residential neighbourhood. When I arrive at the meeting point, I find one of the two staff members sitting on the bench, the Fassino bag upside down, as the other one did. We start. We ring the bell and say, "Hi, we are volunteers for Piero Fassino; can we ask you some questions about the city?" Most people do not respond (also because it's 4 p.m. on a working day). Whoever answers, with rare exceptions, says no. A very strong suspicion withholds them from opening the door. When we find a building door open, we enter and knock on all the doors. We manage to handout three flyers, including one to a couple from the Partito Democratico. The lady is part of the local section and wants to know who we are. She is part of the local section and does not know us (but why wasn't she contacted?). The two staff members are tired and have no enthusiasm. In fact, it takes

> enthusiasm (or a salary) to spend hours ringing doorbells. After about an hour, when we are halfway down the list, the two cannot take it anymore. One of the two jokingly proposes ringing the bells all together. (Fieldwork note, May 13, 2016)

The fact that the staff member did not find anyone available for canvassing shows scarce cooperation with the party. In addition, we see how the lack of involvement in the campaign by the staff members led to an absence of enthusiasm and to consider the time dedicated to the campaign as working time (see the trick of turning the bag inside out once the outing is over). Another problem was the difficulty of collecting data from people who tend to be suspicious, probably because of their fear of fraud or illicit use of their data. Faced with these difficulties, some volunteers openly said they would stop collecting data, thus violating the rules provided by the staff ("I personally collected two email accounts and a phone number, then I stopped asking," Int. 6). Faced with these difficulties and the low number of volunteers mobilized, as the weeks passed, NST tended to become mainly a form of leafleting.

> The staff member told me that today there would be two candidates and two volunteers; instead there are only the two candidates. They are two very young women who were elected in the district in the first round of the elections. They are happy and relaxed, and they do not even handout the flyers to all the people they meet. The activity has been reduced to handing out flyers. The staff member says that it is like this now, that people do not stop anymore, and they are tired. In fact, we walk and many people, especially older people, do not really want to talk to us. And they [the candidates] do not do anything to make them change their minds. (Fieldwork note, June 15, 2016)

Conclusion

Data-driven campaigns are more and more visible in the US and in Europe. In recent years, especially after the Cambridge Analytica scandal, there has been a lot of talk about them. This has led some scholars to underline the gap between rhetoric and practices of data-driven campaigns (Baldwin-Philippi, 2017). Following this stream of literature, in this contribution, I focused on practices, and in particular on how data are actually collected during electoral campaigns. Everyone is talking about social media and algorithms, but we often forget that among the multiple strategies parties use to collect data there is canvassing; that is to say, a direct interaction between a party member, activist, or paid worker and citizens. Some studies have also challenged the dominant view of data-driven campaigns as highly professionalized and sophisticated (Baldwin-Philippi, 2017; Nielsen, 2012).

Canvassing has a long history in the US. And nowadays the interaction between representatives and citizens, what has been called "personalized political communication" (Nielsen, 2012), a feature of the first phase of electoral campaigns, has become popular again. The difference between the interactions of the past and those of modern campaigns is that nowadays data and analytics guide the interaction between campaign volunteers and citizens. However, on the one hand, how, concretely, data is collected by them remains an underdeveloped topic.

The aim of this contribution was to show, through data collected during a direct observation and semi-structured interviews, what is behind data collection in an electoral campaign based on field mobilization and microtargeting. This is particularly interesting because the case of NST represents an adaptation of a technique developed in the US into a different context. Indeed, this was the first campaign of this kind organized in a large Italian city, and we know that different political, social, and institutional contexts can have an impact on the reception of a given campaign tool (Dobber et al., 2017). In particular, studies on the European context are less common than studies of the US (Bhatti et al., 2019).

Our results point out that, despite the campaign aiming to recruit volunteers among lay citizens, most volunteers are indeed party members. If, on the one hand, this could mean party members continue to have an important role in the European context, mainly as organizational resources and as a source of legitimacy (van Haute and Gauja, 2015), it is also true that, through the representations of the campaign of party members, the difficulties in the rooting and mobilization of the party emerged. Contextual differences between Europe and the US made it difficult to accept for citizens a role of "political volunteer." This can be considered one of the reasons for the scarce appeal of the campaign for lay citizens.

Regarding data collection, this study confirms the result, presented in similar studies, that volunteers and activists have room to manoeuvre in the interpretation of their role and that, even when the organization's strategy is highly rationalized, the actual practices implemented by militants and volunteers can differ from it and lead them to diverge from instructions (Belkacem and Talpin, 2014; Nielsen, 2012). If then for some participants (staff members), the ones most aware of the mechanism behind data collection, the collection of data becomes an end in itself, others (volunteers/party members) do not understand the reason behind it and soon stop asking citizens for their data. These resistances could depend both on the different context, as we have seen, and on members' resistances to practices they are not used to and that they perceive as an imposition from above (Lefebvre, 2016).

To sum up, the account of this campaign shows that rhetoric and practices of data-driven campaigns diverge and confirm the assumption that these campaigns are often less professionalized and sophisticated than one may imagine. In a context in which data informs parties' strategies, knowing how data is collected and the potential pitfalls of the data collection can help us to understand and put into perspective the risks of these practices for democracy. From this perspective, microtargeting appears less harmful than it looks.

Notes

[1] Gerber and Green's experimental research shows that door-to-door activity is more effective than posters, emails, and phone calls. In fact, canvassing allows parties to "earn" one vote for every fourteen conversations, against one for every thirty-eight telephone calls and one for every hundred thousand flyers.

[2] In the United Kingdom, canvassing is a consolidated practice because of the incentives of the uninominal electoral system. Even in this country we have witnessed, especially from 2015 onwards, the profound Americanization of electoral campaigns, with both the Conservatives and Labour using US political consultants who previously worked for Obama.

[3] As regards the use of the word "contagion," Robert (2007, p. 16) invites us not to think about the importation of managerial practices in politics in terms of simple

diffusion, in which the new practices do not face resistance or undergo transformations. In his opinion, it is more appropriate to investigate the strategic appropriation of these practices by actors and the reasons for their interest in them.

[4] See Giambartolomei, A. *Fassino vuole fare Obama: studenti arruolati a Torino*, in "Il Fatto Quotidiano," June 1, 2016.

[5] In the first phase, candidacies are submitted to members in a "one member one vote" procedure. The first three candidates are admitted to the second phase if they have obtained at least 5 percent of the votes and at least 15 percent in five regions.

References

Almond, G., and Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton University Press.

Altheide, D.L., and Snow, R.P. (1979). *Media logic*. Sage.

Baldwin-Philippi, J. (2017). The myths of data-driven campaigning. *Political Communication*, *34*(4), 627–33.

Bennett, C.J., and Lyon, D. (2019). Data-driven elections: Implications and challenges for democratic societies. *Internet policy review*, *8*(4).

Bhatti, Y., Dahlgaard, J.O., Hansen, J.H., and Hansen, K.M. (2019). Is door-to-door canvassing effective in Europe? Evidence from a meta-study across six European countries. *British Journal of Political Science*, *49*(1), 279–90.

Biorcio, R., Caruso, L., and Vitale, T. (2016). Le trasformazioni del sistema politico italiano e l'associazionismo. In R. Biorcio and T. Vitale (Eds.), *Italia civile* (19–30). Donzelli.

Blumler, J.G. (2016). The fourth age of political communication. *Political Communication*, (1), 19–30.

Blumler, J.G., and Kavanagh, D. (1999). The third age of political communication: Influences and features. *Political Communication*, *16*(3), 209–30.

Bodó, B., Helberger, N., and de Vreese, C.H. (2017). Political micro-targeting: A Manchurian candidate or just a dark horse? *Internet Policy Review*, *6*(4), 1–13.

Castells, M. (2009). *Communication power*. Oxford University Press.

Cepernich, C. (2017). *Le campagne elettorali al tempo della networked politics*. Laterza.

Dalton, R.J., and Wattenberg, M.P. (Eds.). (2002). *Parties without partisans: Political change in advanced industrial democracies*. Oxford University Press.

Dobber, T., Trilling, D., Helberger, N., and De Vreese, C.H. (2017). Two crates of beer and 40 pizzas: The adoption of innovative political behavioural targeting techniques. *Internet Policy Review*, *6*(4), 1–25.

Dommett, K. (2019). Data-driven political campaigns in practice: Understanding and regulating diverse data-driven campaigns. *Internet Policy Review*, *8*(4).

Easton, D. (1965). *A framework for political analysis*. Prentice-Hall.

Flinders, M., and Wood, M. (2014). Depoliticisation, governance and the state. *Policy and Politics*, *42*(2), 135–49.

Gerber, A.S., and Green, D.P. (2000). The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *American Political Science Review*, *94*(3), 653–63.

Gerber, A.S., and Green, D.P. (2004). *Get out the vote: How to increase voter turnout.* Brookings Institution.

Gorton, W.A. (2016). Manipulating citizens: How political campaigns' use of behavioral social science harms democracy. *New Political Science*, *38*(1), 61–80.

Grilli di Cortona, P. (2007). *Il cambiamento politico in Italia: dalla prima alla seconda Repubblica*. Carocci.

Lefebvre, R. (2016). La modernisation du porte-à-porte au Parti socialiste. Réinvention d'un répertoire de campagne et inerties militantes. *Politix*, *29*(1), 91–115.

Liégey, G., Muller, A., and Pons, V. (2013). *Porte à porte. Reconquérir la démocratie sur le terrain*. Calmann-Lévy.

Lipset, S.M., and Rokkan, S. (1967). *Party systems and voter alignments: Cross-national perspectives.* Free Press.

Manin, B. (1995). *Principes du gouvernement représentatif.* Calmann-Lévy.

Mazzoleni, G., and Schulz, W. (1999). "Mediatization" of politics: A challenge for democracy? *Political communication*, *16*(3), 247–61.

Michels, R. (1911). *Zur Soziologie des Parteiwesens in der modernen Demokratie*. Werner Klinkhardt.

Moro, G. (2013). *Cittadinanza attiva e qualità della democrazia*. Carocci.

Natale, P. and Fasano, L. (2017). *L'ultimo partito: 10 anni di Partito democratico*. Giappichelli.

Nielsen, R.K. (2012). *Ground wars*. Princeton University Press.

Norris, P. (2000). *A virtuous circle: Political communications in postindustrial societies*. Cambridge University Press.

Robert, C. (2007). Les transformations managériales des activités politiques. *Politix*, (3), 7–23.

Scarrow, S. (2014). *Beyond party members: Changing approaches to partisan mobilization*. Oxford University Press.

Talpin, J., and Belkacem, R. (2014). Frapper aux portes pour gagner les élections? Ethnographie de la campagne présidentielle socialiste dans deux villes du Nord de la France. *Politix*, *27*(1), 185–211.

Van Biezen, I., Mair, P., and Poguntke, T. (2012). Going, going, … gone? The decline of party membership in contemporary Europe. *European Journal of Political Research*, *51*(1), 24–56.

van Haute, E. and Gauja, A. (Eds.). (2015). *Party members and activists*. Routledge.

van Zoonen, L. (2005). *Entertaining the citizen: When politics and popular culture converge*. Rowman and Littlefield.

Vassallo, S., and Passarelli, G. (2016). Centre-left Prime Ministerial Primaries in Italy: The laboratory of the 'open party' model. *Contemporary Italian Politics*, *8*(1), 12–23.

Wood, M., and Flinders, M. (2014). Rethinking depoliticisation: Beyond the governmental. *Policy and politics*, *42*(2), 151–70.

Zuiderveen Borgesius, F., Möller, J., Kruikemeier, S., Ó Fathaigh, R., Irion, K., Dobber, T., Balazs, B., and de Vreese, C.H. (2018). Online political micro-targeting: Promises and threats for democracy. *Utrecht Law Review*, *14*(1), 82–96.

This page intentionally left blank

PART 3

Policy: Surveillance and Data Protection during the Coronavirus Pandemic

This page intentionally left blank

7

Surveillance Capitalism Meets the Pandemic

Surveillance Challenges to the "Social Contract"

David Lyon

Introduction

Data-dependent "solutions" to problems posed by the COVID-19 pandemic gave surveillance capitalism,[1] along with willing governments, further opportunities to develop data-focused modes of addressing social and political problems. In so doing, they frequently further reduced both freedom and fairness in the name of extraordinary emergency measures. This conjunction strengthened public-private partnerships and simultaneously highlighted the need for a contemporary reset of any notion of a social contract insofar as that idea has salience for today's "data democracy."[2] Increasingly, it seems, government-and-business partnerships, on the one hand, and digitally disempowered citizens and consumers, on the other, struggle to recognize each other, let alone to develop a meaningful modus vivendi for a democratic global and planetary future. This chapter is about policy, but it emphasizes the role of civil society in particular as a vital counterweight to the combined power of government *and* corporation.

The argument that follows has two main parts. The first is on the challenges of data democracy that are presented by the recent meeting of surveillance capitalism and the pandemic. The evolving relationships between "platform-and-pandemic" raise more issues than the familiar "function creep" practices. Platforms are entering more spheres, with new consequences. Think of the Google-Apple Exposure Notification (GAEN) Application Programming Interface (API) to enable digital contact-tracing apps, for example. This hugely bolstered the position of these tech giants in the public health domain and offers a key example of expanding infrastructural surveillance power (Solano et al 2022).

Specific surveillance technologies are also gaining traction despite evidence of their limited usefulness, well-known threats to privacy, and debatable levels of success. As Mark Andrejevic and Neil Selwyn (2022) say of facial recognition technologies, "The experience of the past couple of decades has been shaped by the widespread implementation of increasingly comprehensive and granular forms of monitoring in exchange for the convenience and affordances of various data-driven digital technologies" (viii).

The language of trade-off and bargains features frequently in such debates, and that is why the notion of "social contract" is arguably germane to current debates over surveillance capitalism and the COVID-19 pandemic (Dans, 2020). So, the second part of my argument concerns the prospects for the relatively recent revival of social contract approaches to the issues raised by so-called big data surveillance, algorithmic analysis, artificial intelligence, and machine learning. The early modern notion of a social contract, revived by Rawls in the 1970s, now has new life in the hands of current observers across several disciplines: political science, psychology, sociology, computer science, and business studies (see, for example, Al-Rodhan, 2014; Chesterman, 2011; Kruikemeier, Boerman, and Bol 2020; Liaropoulos, 2020; Martin, 2015; Pallitto, 2020; Rahwan, 2018). How conducive is this to fostering new approaches to data democracy?

Vaccine Passports as Surveillance

Vaccine passports, issued around the world as the pandemic was met by widespread vaccination in wealthier countries, are but one example of "pandemic surveillance." They exemplify the issues in that not just matters of "privacy" but also of "data injustice" become apparent. Quebec, the first Canadian province to approve vaccine passports, was immediately roiled with controversy at their introduction. Bar owners complained that instead of greeting patrons, they had to interrogate them.

An issue erupted in the National Assembly over whether "les deputés" (MNAs) had to show their passports – were they "essential workers" or not? Meanwhile the Ligue des droits et libertés ("League of Rights and Liberties") worried about data security and the possibility that the passport could be used for other monitoring purposes.

Australia, similarly to Canada, had a policy of state decision making about vaccine passports. Apart from federal rules about proof of vaccination at international airports, each jurisdiction made its own rules about proof of vaccination (Karp, 2021). Andrew Barr, chief minister of the Australian Capital Territory, said that the territory would not issue vaccine passports for human rights and practical enforcement reasons. The reasoning seemed to be: Why should vaccinated people be privileged with extra freedoms, while others, who for whatever reason remain unvaccinated, be denied services or access (see French and Monahan, 2020)?

Not for the first time during the pandemic, then, the spectre of surveillance surfaced, this time with vaccine passports. Pressure to roll them out came from travel and hospitality companies anxious for a return to business as usual, and newscasts boasted constantly of their success in reducing COVID-19 infection rates. But amid the various arguments about such passports, a further conversation was largely absent – surging surveillance enabled by yet another pandemic innovation (see Lyon, 2022).

Of course, the pull of "back to normal" was understandably strong after nearly two years of tragic death rates, anxieties over personal and family health, and constant lockdowns and other restrictions. But the calls for a "just recovery" have often fallen on deaf ears, especially with respect to the massive uptick in monitoring and surveillance that puts those following 9/11 in the shade. Those demanding a just recovery demand that the uneven impacts of the pandemic – especially along the lines of class, race, and gender – be addressed in concrete ways that are manifestly fair.

So, what is the connection between surveillance and *injustice*? Is not *privacy* the core issue? Well, privacy *is* challenged in new ways, perhaps above all in the domestic context, which is targeted in unprecedented fashion during the pandemic's oft-repeated "stay home, stay safe" mantra. But the privacy questions are not identical for all. If you work remotely, there is a vast difference between the experience of comfortable white-collar professionals and precarious gig workers, for example. Also, the "stay home" rule increased the burden on women, including their often-unwelcome exposure on screen to unknown others, while working, shopping, or learning online. Privacy, so often considered in *individual* ways, clearly affects *groups* of people, as the pandemic profoundly illustrates (Puri, 2021, Austin, 2022).

As in other areas, pandemic surveillance is both experienced more profoundly among those already disadvantaged (Toh and Brown, 2020) *and* often serves to retrench rather than rectify such disadvantages. This is certainly true for Indigenous Peoples in North America, who suffer from both insufficient surveillance (their condition frequently lacks adequate data) *and* slanted surveillance (mistaken assumptions are made by health authorities about their pandemic knowledge and practices) (Hendl and Roxanne, 2022). And the vaccine passports are indelibly surveillant – they rely on government-held personal health data, which citizens are obliged to display for travel or social participation. They make their holders visible to airlines, restaurants, sporting events, and the like, representing them as "responsible citizens" and permitting movement or access to those holders.

As it happens, there are many valid reasons why someone might not hold a vaccine passport – they may be immuno-challenged, for instance, or fear that the warning labels on Pfizer and Moderna about the possibilities of heart inflammation may affect them as seniors with a history of cardiac problems. They may also have read the small print on vaccine limits or side effects and decided to wait for improved options. But the choices carry consequences. As well as being refused boarding at an airport or being served in a bar, holders may be scapegoated. They may fill a new role as an excludable, risky other.

Controversy over vaccine passports, due to their potential divisiveness, was predicted early on by the World Health Organization. The report of the Ada Lovelace Institute to the UK government warns about this (Parker, Montgomery, and Freeguard, 2021), as does the Office of the Privacy Commissioner in Canada (OPC, 2021). The Waterloo, Ontario, based Centre for International Government Innovation observes that the expansion of digital IDs in the guise of such passports will entail a likely control creep outcome (Renieris, 2021). Also in the UK, public health researcher Robert Dingwall (2021) decried the arrival of the "biosecurity state" in vaccine passports. One repeated concern from many citizens was how and when these may be phased out. By the end of February 2022, such passports were no longer needed for pubs, restaurants, and theatres and, from April 1, 2022, the government recommended that the National Health Service (NHS) COVID Pass be used only for foreign travel (which means that the infrastructure enabling them has remained in place).

The Wider Picture

In fact, many kinds of surveillance, introduced for the pandemic, could also be here to stay. But are they being questioned now, as the possibilities for a postpandemic future emerge? Not so much. But red lights began to flash almost as soon as COVID-19 was identified, when a raft of new surveillance techniques – from digital contact-tracing onward – were hastily rolled out courtesy of over-eager tech companies and underprepared public health authorities (Kitchin, 2021).

Big data "solutions" were front and centre, offering the modelling behind the now-familiar dashboards of pandemic progress and the means to remotely track and trace potentially infected people. Of course, there is value in these initiatives; some have been credited with saving lives or preventing ICUs from being overwhelmed. But in Ontario, the Personal Health Information Protection Act was quickly and quietly altered within an omnibus bill to facilitate increased data access (Scassa, 2021). And as it transpired, most Canadians never used the COVID-19 Alert app, partly because of privacy concerns and accessibility and inclusion challenges, affecting vulnerable populations such as seniors (Phillips and Mamuji, 2021).

At the same time, the pandemic prompted massive surveillance expansion in other areas of social life, especially via the stay home, stay safe mantra. Working, learning, shopping, and being entertained at home also entail an explosion of domestic surveillance, through platforms such as Prodoscore for monitoring employee performance, Zoom for myriad online conferencing purposes, Examity for remote proctoring of tests and examinations, Amazon for universal shopping, plus movie and music providers to compensate for the erosion of public entertainment – and public life in general. Tech companies pivoted from their previous primary activities to others, not only to public health but also to other pandemic-related areas including policing. Every platform that repurposed itself for the pandemic is highly surveillant (Lyon, 2022).

Surveillance Capitalism Meets Pandemic

In each case, whether in public health or stay-home initiatives, surveillance capitalism – and tech companies in general – are also implicated, especially through public-private partnerships between government departments and platforms. Indeed, some speak of an "epidemiological turn in digital surveillance" with two dimensions, "function creep and market-making" (Taylor et al., 2020). On the one hand, already-existing systems were upgraded or repurposed for the COVID situation – sometimes developed by mobile network operators for low-to-middle income countries over recent decades. And on the other, software developers launched mobile apps to support contact tracing, often based on the GAEN API. Such developers believe that they too can benefit from the use of the apps – Google has been seeking greater means of access to health data for years (Lyon, 2022) – which is where surveillance capitalism becomes evident.

Surveillance capitalism, whose emergence was noted by people such as Vincent Mosco (2014) and John Foster and Robert McChesney (2014), is now associated above all with the name of Shoshana Zuboff (2015) and her 2019 blockbuster, *The Age of Surveillance Capitalism.* In it, she describes how value is extracted from data created as a by-product of the everyday use of digital technologies – primarily the smartphone. Her analysis of Google's role in this process is especially telling, even if she depends quite heavily on their spokespersons for the analysis.

Zuboff argues that surveillance capitalism bears strong resemblance to B.F. Skinner's *Walden Two,* a behaviourist dystopia where technology drives decision making and compensates for the deficiencies of human decision making: free will and autonomy. This system uses data to predict outcomes and to produce addiction among users that, together, offer unprecedented opportunities for platform companies to manipulate and exploit the inner lives of their users.

As Kirstie Ball (2019) pointed out in her review of *The Age of Surveillance Capitalism*, Zuboff also – helpfully – makes much use of the language of "securitization," often referring to the "states of exception" and "rendition" visible in the activities of surveillance capitalism today. The suspension of normal democratic rules (exception) and the use of coercion (rendition) have been debated by Giorgio Agamben (2005) and others, above all since 9/11. Curiously, these military-security metaphors for the "war" against terrorism were revived during the COVID-19 pandemic, aided by the same surveillance capitalist platforms, not to mention actual deployment of military personnel and technology in this "struggle." Such may easily undermine social contract ideas, as François Pellegrini notes in this volume.

Zuboff (2021) has also written about the role of the pandemic in boosting surveillance capitalism, in terms consistent with a "social contract" reset:

> The United States and the world's other liberal democracies have thus far failed to construct a coherent political vision of a digital century that advances democratic values, principles and government. While the Chinese have designed and deployed digital technologies to advance their system of authoritarian rule, the West has remained compromised and ambivalent.

Journalists, too, have made these connections; for instance, Victoria Kim (2021) reported from South Korea on the repurposing of "smart city" technologies for tracking the spread of COVID in urban areas. She, too, notes the challenge of such developments to implicit commitments to a "social contract."

Social Contract Reset?

Later twentieth-century surveillance developments, based on "information," were met with updates to the "social contract" (Weller, 2012). Today, an increasing number of voices argue that the social contract should once again be revamped due to the rapid growth of data analytics and the rise of surveillance capitalism. No doubt many will ask what it will take, given the vast new surveillance powers that have been unleashed since social media, platform infrastructures, and especially surveillance capitalism emerged.

In what follows, we shall critically examine some contributions to the surveillance-and-social-contract debates, especially their treatment of the civil society dimension of the "social contract," on the one hand, and the demand for government to take a strong and principled role in holding tech giants to account, on the other.

The notion of social contract, derived originally from early modern thinkers, including Hobbes, Locke, and Rousseau, has been adapted, in the face of social, economic, and political change, for the past three hundred years. Notably, the security, protection, and welfare gradually offered by governments in return for the loss of sovereignty among citizens has increasingly revolved around issues of information and, laterally, data. In everything from conscription-related data for the call-up of soldiers for war to the data amassed for the purpose of welfare provision, the social contract has been under question and gradually – sometimes painfully – revised (Higgs, 2004, 150).

From the nineteenth century, bureaucratic monitoring arose in relation to industrialization, using new technologies – typewriters, telephones, cameras, and so on – to govern citizens as "things" might be governed in emerging industrial processes. Although this is less frequently noted, the monitoring of *consumers*, using new technologies, also expanded from the later nineteenth century (Lauer, 2017). This process intensified during the twentieth century, as did both warfare and welfare – with the rise of state security intelligence (Jeffreys-Jones, 2017) and the welfare state (Garland, 2016), not to mention the later twentieth-century growth of targeted advertising (Gandy, 2021), which "required" increasing amounts of personal data that were processed in increasingly arcane ways.

Thus Toni Weller (2012, 61) suggests that today's variant of the classic social contract in democratic states is that "citizens accept surveillance in order to ensure that they are entitled to welfare and protected from threat as long as the means of such surveillance are transparent and accountable." However, Weller was writing between the attacks of 9/11 and the revelations about National Security Agency (NSA) surveillance leaked by Edward Snowden during the first decade of social media. Since then, the stakes have been raised considerably with the development of monopolistic platform companies (Clement and Lyon, 2018) – sometimes referred to collectively as "platform capitalism" (Srnicek, 2016) – that operate by establishing hardware and software systems that are actually used by other companies building on the "platform." Platform companies supply the digital infrastructure for others – such as Instagram, Twitter, YouTube, Uber, and Airbnb.

Controversies about how far this "new variant" is really working have become more and more strident in the twenty-first century, in North America especially after 9/11 and then increasingly worldwide with the rise of social media – seen sharply in the Cambridge Analytica–Facebook debacle – the Snowden revelations in 2013, and the COVID-19 pandemic in 2020–23. All too frequently, so far from actively consenting to surveillance, citizens and consumers are *unaware* that they are under surveillance and especially ignorant of how surveillance might affect them. And at the same time, questions of government or corporate transparency can easily be used to *deflect* attention from the deeper issues of accountability.

Interestingly, with the rise of China as a leading player in the digital world, the questions find a new foil in the growth of authoritarian capitalism and its associated surveillance – much criticized in the Western world. But this also conveniently deflects attention from what is *actually happening* in supposedly social-democratic polities, where surveillance power grows especially through public-private partnerships, seen in bold relief during the pandemic. Whether in China or in liberal democracies, however, the pandemic must be understood as an opportunity for the largest-ever global surveillance surge to appear.

Debating Surveillance and Social Contract

Several insightful and thoughtful contributions have been made in recent years that apply ideas and practices of social contract to the growth of surveillance. These apply both to state surveillance – and especially to national security intelligence – and to consumer surveillance in an era of surveillance capitalism. None, as yet, speaks to the conjunction of these two, as revealed in the pandemic, where techniques and technologies from consumer surveillance are used in the service of government mandates such as mask wearing, quarantine observance, contact tracing, or physical location monitoring.

Simon Chesterman (2011), for example, takes a cue from Edward Shils (1956), who argued that liberal democracy rests on protecting privacy for individuals and rejecting it for government. Chesterman describes the ways that the US has rapidly taken the opposite tack, arguing that an increasingly urgent need is for new levels of government accountability. His case is that growing international threats, the spread of digital technologies, and the emerging culture of data sharing prompted an implicit social contract in which the state and other actors have power over data in exchange for the security and convenience of living in the modern world. However, an alternative model of a workable social contract is not really explained by Chesterman.

More recently, Robert Pallitto (2020) has argued for a new social contract under the title *Bargaining with the Machine*. He is concerned for privacy but also for justice, given the inequalities associated with and exacerbated by current surveillance trends. It is about how people negotiate trade-offs in the face of "growing structural impediments" to any semblance of free choices in this context (104). Pallitto acknowledges the weakness of the post-9/11 "liberty-for-security" trade-off, arguing instead for simultaneous structural change affecting state-corporate actors *and* microlevel decision making by individual agents.

On the civil society side, Pallitto (2020) maintains that considerations of social values, externalities, and protecting dignity are all significant. Cost-benefit analysis does not cut it for Pallitto; "social values" refer to what is important in human relationships, for example. Externalities are the negative impacts of consumer choices (e.g., the massive energy consumption of internet and its data servers [Griffiths, 2020]). One may seek privacy as care for others who may be harmed by our failure or to protect user dignity, for instance, through caring about personal medical records.

These themes are also picked up in recent "social contract" writings, some following the controversies regarding data use by national security agencies after the revelations by Edward Snowden and others prompted by the pandemic. Andrew Liaropoulos (2020), for instance, argues that corporate power began to seriously threaten the social contract from the late twentieth century, and the turn to "big data," or what he calls the "state of data," is exacerbating this in the twenty-first. This demands new trust-building efforts involving citizens, *corporations*, and the state, with a view to redirecting efforts to meeting social need. With new market regulation and accountability, plus trust building, there is, he argues, hope for a new "digital social contract."

Similarly, Nayef Al-Rodhan (2014) begins by commenting on the civil liberties threat posed by the NSA's use of metadata, which he describes as mass, unaccountable surveillance that undermines governmental power, especially the rule of law. But he also observes corporate power at work here, especially through "monopoly platforms." The social contract, in this view, is essential to ensure that the state has the interests of the people at heart, which means restraining corporate power and reinstating the rule of law. If data analytics is to serve the common good, then upholding civil liberties *and* seeking corporate and government accountability is essential.

Others situate their case for a revived social contract within specific domains, rather than in blanket – societal – terms. Iyad Rahwan (2018), for instance, speaks from within the domain of data analysis, arguing for an "algorithmic social contract" to ensure that artificial intelligence is fair and accountable. He understands a strategy of extending humans-in-the-loop practices to "society-in-the-loop" as a means of negotiating the values of stakeholders. This would, in his view, enable human and societal supervision of AI, for instance, in developing smart highways, thus "taming the techno-leviathan." Similar but sharper arguments are made by Joy Buolamwini (2023).

Within business studies, Kirsten Martin (2015) offers some telling insights on how the terms of the data-oriented social contract must shift from obtaining the consent of the user, consumer, or employee to making it the responsibility of the firm as a contractor to "maintain a mutually beneficial and sustainable solution." Building on Helen Nissenbaum's (2009) work, she argues that users should be able to discriminately share information within situations of "contextual privacy." Facebook (now Meta) or Verizon should give users the right to know who has access to their data and how it us used. People should have *space* (Bennett, 1992) for freedom of movement and access to data and to how it is used within a relationship. Expecting users to read and understand privacy policies must stop – the social contract makes for relationship and "strong community."

The story is picked up within media psychology by Sanne Kruikemeier, Sophie Boerman, and Nadine Bol (2019) who demonstrate the need for a new social contract from their empirical studies of users. From their research, they perceive that users do have a vague sense of a contract *and* a sense that it is being breached. About 50 percent of their respondents (in the Netherlands) think that the social contract is unreliable and only about 4.75 percent believe that it is reliable.

There is thus an apparently rising interest in how notions of social contract might be applied to the world of data analytics, AI, monopolistic platforms, and surveillance capitalism. And so, it is worth standing back and evaluating the potential for using this concept more broadly within debates and practical policy responses to this rapidly expanding milieu that is visible at scale, from local to global.

Developing a Social Contract for a Digital Society

Given that social contract theory covers a multitude of possibilities, let me comment briefly on some advantages of using the concept to catalyze a reset of relationships within a digital society. First, classical social contract theory has in some versions tended to be individualistic, but this is not a necessary aspect. Understanding how, culturally, consumers and users relate to the digital, openings for civil society are clearly present as a feature, for example, of the very groups enabled by social media and many platforms.

Second, a social contract suited to today's digital society, where surveillance capitalism is prominent, is obliged to consider the three-way relationships between civil society, on the one hand, and the state *and* corporation, on the other. This means moving beyond mere calculation and mutual benefit to developing relationships that are fair, which is where notions such as data justice (Dencik and Sanchez-Monedero, 2022; Taylor, 2017) come in. This obliges all parties to negotiate appropriate policy that will be underwritten by legal requirements; for example, excluding monopoly, "weblining," and so on.

Third, we must also ensure that the language of the social contract goes beyond law, limits, and policy. The point of seeking greater transparency about data collection and use, for instance, is to ensure accountability, on the one hand, and trust, on the other. The quality of relationships that are vital for the social contract for digital societies cannot ultimately depend on legal and calculative measures alone. They have broader aims in view to foster accountability and trust and such aspirations are conducive to the *quality* of relationships that could characterize such digital social contracts.

Democratic Challenges for the Digital Social Contract

So, what role might the popular myth – or "imaginary" (Taylor, 2003) – of the "social contract" play, postpandemic? How can the notion of a fair deal between state and citizen be restored – or, perhaps reframed or rediscovered – after the pandemic? Clearly, the role of surveillance capitalism has been prominent during the pandemic, often seen in key developments such as the Apple-Google API for contact-tracing apps.

There are many structural barriers to enhancing digital citizenship in a datafied society (Hintz, Dencik, and Wahl-Jorgensen, 2019, 152). One key issue is the way that relations of power translate into influence over public debates about data collection, analysis, and use. People may be alerted to the realities of rapidly rising surveillance but counter-narratives about security, and now public health, quickly and strategically arise to tamp down negative critique. Dreams of digitally enhanced citizenship patterns are punctured by the realities of surveillance capitalism, in which the amassing, analyzing, and use of data by corporations and governments lead to more authoritarian practices.

Data ownership and its associated concentrations of wealth and power in a small number of technology monopolies and other platform companies is a key issue militating against the growth of digital citizenship of more meaningful kinds. Yet there are also groups arguing more strenuously for participatory practices in the design and use of digital infrastructure and finding ways to develop a "digital commons" where shared responsibilities are paramount.

I have argued for some time that the ways forward for digital citizenship have to be based on the ways ordinary internet users may show themselves to be more than the "things" that James Beniger argued they began to be treated as in the twentieth century (or as Deleuze [1992] saw them – "dividuals" or mere data clusters). For, as Ian Hacking (2006) argues, there exists what he calls a "looping effect" in which those under surveillance become aware of the surveillance and its effects, even if only in a mild way, as they feel spooked by the speed with which advertising appears in their feed after they have written a message or ordered some merchandise online. But that "looping" shows how those affected may be *treated* like things but may not be *reduced* to things. In that looping process, we see a crack that lets in the light – light that may emerge as democratic opportunity (see also Lyon, 2018).

Simultaneously, other features are required on the landscape of any renewed social contract. For one thing, there should be clarity about what sorts of issues have to be confronted and in what priority order (see, for example, Deibert, 2020). One key area relates to the ways issues of surveillance have been questioned over the past decades, indeed, since the late nineteenth century – under the rubric of privacy. Now, privacy has unavoidably *social* dimensions, but the ways everyday internet users appeal to the concept shows that many think only of the *individual* aspects of privacy – seen as a personal as opposed to a social matter. We will have to live with privacy laws and their related international organizations for a while, but to address the issues raised here, more is needed.

"Data justice" (Dencik and Sanchez, 2022; Taylor, 2017) is a concept that resonates with the concerns about inequality and injustice perpetrated by contemporary surveillance. Today's surveillance practices are inextricably tied up with modes of scoring and ranking consumers, citizens, and indeed all exposed to today's platforms. This is the world exposed – helpfully – to the reality of data *injustice* through television series such as *Black Mirror*. Data justice reminds us that surveillance uses data to make ordinary people visible to those in power, whether government, corporation, or a combination. But it also *represents* them according to its criteria and *treats them accordingly*. Vaccine passports, where this article began, make visible those without vaccination, representing them as irresponsible or careless or worse, and allows them to be treated negatively, by exclusion from privileges or needed services.

But even if one is equipped with some strong notion of data justice, there still must be engagement in order to act as digital citizens. Pallitto (2020) starts a discussion of this. He observes, tellingly, that consumer desire reproduces consent to a consumer society. He points out that instead of simply "inviting surveillance" by our everyday online interactions, we should rather find ways to express social values, be aware of "externalities," and see protecting personal dignity as an obligation. He maintains that we should all take our online interactions more seriously. But this kind of work requires fuller analysis and debate – and also work that is more fully attuned to the already-existing questions about surveillance and "social contract."

At the end of the day, the question remains: How far does a "social contract" really exist in relation to contemporary surveillance in a world increasingly dominated by the powerful platforms of surveillance capitalism (Kruikemeier, Boerman, and Bol, 2019)? That the idea could, and I venture to say *should*, be explored is an important and timely one. But it will take much careful research to draw out the key issues, which here I have argued should be couched as "data justice" along with privacy or data protection. It will also involve creating a coalition of forces that oppose surveillance on a general level and those concerned especially about the negative effects of surveillance on marginalized groups (Franks, 2017).

And it will involve the "conscientization" – a critical awareness of one's social reality through reflection and action (Freire, 1970) – of internet users – all of us – to be aware of more issues than just *our own* dignity or internet "freedom." It will require new content in terms such as "state" (when corporate power is so deeply inscribed in government activity) and "civil society" (when so frequently the latter is seen merely as "voluntary associations" and not as vital players in determining the terms of the social contract). But now is a crucial moment for such radicalizing debates to occur (see Rone and Biancalana, this volume) – when the pandemic has unleashed the largest surveillance surge ever and few hesitated to ask whether the results of this tech-solutionist surge were either fit-for-purpose, privacy-protective, or supportive of data justice.

Notes

[1] "Surveillance" is understood in this chapter in specific ways. See Lyon 2022b.

[2] I take it that "data democracy" encourages technology for the common good and protects against invasive, discriminatory, and exploitative uses; supports platform accountability and user rights, limits collection and use of personal data; reduces government censorship and rejects unwarranted surveillance both governmental and commercial. These kinds of purposes are behind, for example, the Centre for Technology and Democracy: https://cdt.org/.

References

Agamben, G. (2005). *The state of exception*. University of Chicago Press.

Al-Rodhan, N. (2014). The social contract 2.0: Big data and the need to guarantee privacy and civil liberties. *Harvard International Review*, 16.

Andrejevic, M., and Selwyn, N. (2022). *Facial recognition*. John Wiley and Sons.

Austin, L. (2022). From Privacy to Social Legibility. *Surveillance & Society 20*(3), 302–5.

Ball, K. (2019). Review of Zuboff's *The age of surveillance capitalism*. *Surveillance and Society*, *17*(1/2), 252–56.

Bennett, C.J. (1992). *Regulating privacy: Data protection and public policy in Europe and the United States*. Cornell University Press.

Buolamwini, J. (2023). *Unmasking AI*. Penguin.

Chesterman, S. (2011). *One nation under surveillance: A new social contract to defend freedom without sacrificing liberty*. Oxford University Press.

Clement, A., and Lyon, D. (2018, April 23). Facebook: A mass-media, micro-surveillance monopoly. *The Globe and Mail*. https://www.theglobeandmail.com/opinion/article-facebook-a-mass-media-micro-surveillance-monopoly/.

Dans, E. (2020, March 23). Coronavirus, surveillance and the redefinition of the social contract. *Forbes Magazine*. https://www.forbes.com/sites/enriquedans/2020/03/23/coronavirus-surveillance-and-the-redefinition-of-the-socialcontract/.

Deibert, R.J. (2020). *Reset: Reclaiming the internet for civil society*. House of Anansi Press.

Deleuze, G. (1992). Postscript on the societies of control. *October 59* (Winter), 3–7.

Dencik, L., and Sanchez-Monedero, J. (2022). Data justice. *Internet Policy Review*, *11*(1).

Dingwall, R. (2021, April 29). Resisting the biosecurity state. *Social Science Space*.

Foster, J.B., and McChesney, R.W. (2014). Surveillance capitalism: Monopoly-finance capital, the military-industrial complex, and the digital age. *Monthly review*, *66*(3), 1.

Franks, M.A. (2017). Democratic surveillance. *Harvard Journal of Law and Technology*, *30*(2), 426–89.

Freire, P. (1970). *A pedagogy of the oppressed*. Continuum.

French, M., and Monahan, T. (2020). Dis-ease surveillance: How might surveillance studies address COVID-19?. *Surveillance and Society*, *18*(1), 1–11.

Gandy, O. (2021). *The panoptic sort: A political economy of personal information*. Oxford University Press.

Garland, D. (2016). *The welfare state: A very short introduction*. Oxford University Press.

Griffiths, S. (2020, March 6). Why your internet habits are not as clean as you think. *BBC*. https://www.bbc.com/future/article/20200305-why-your-internet-habits-are-not-as-clean-as-you-think.

Hacking, I. (2006). Making up people. *London Review of Books*, *28*(16).

Hendl, T., and Roxanne, T. (2022). Digital surveillance in a pandemic response: What bioethics ought to learn from Indigenous perspectives. *Bioethics*, *36*(3), 305–12.

Higgs, E. (2004). *The information state in England*. Palgrave-Macmillan.

Hintz, A., Dencik, L., and Wahl-Jorgensen, K. (2019). *Digital citizens in a datafied society*. Polity Press.

Jeffreys-Jones, R. (2017). *We know all about you: The story of surveillance in Britain and America*. Oxford University Press.

Karp, P. (2021, September 24). Vaccine passports in Australia. *The Guardian*. https://www.theguardian.com/australia-news/2021/sep/25/vaccine-passports-in-australia-who-will-impose-them-and-how-will-they-work.

Kim, V. (2021, December 9). Who's watching? How governments used the pandemic to normalize surveillance. *LA Times*. https://www.latimes.com/world-nation/story/2021-12-09/the-pandemic-brought-heightened-surveillance-to-save-lives-is-it-here-to-stay.

Kitchin, R. (2021) Civil liberties or public health, or civil liberties and public health? Using surveillance technologies to tackle the spread of COVID-19. *Space and Polity*. https://doi.org/10.1080/13562576.2020.1770587.

Kruikemeier, S., Boerman, S.C., and Bol, N. (2019). Breaching the contract? Using social contract theory to explain individuals' online behavior to safeguard privacy. *Media Psychology*, *23*(2), 269–92.

Lauer, J. (2017). *Creditworthy: A history of consumer surveillance and financial identity in America*. Columbia University Press.

Leaders. (2020). "The state in the time of COVID-19," *The Economist,* March 26.

Liaropoulos, A. (2020). A social contract for cyberspace. *Journal of Information Warfare*, *19*(2), 1–11.

Lyon, D. (2018). *The culture of surveillance*. Polity Press.

Lyon, D. (2022a, May 18). *Beyond big data surveillance: Freedom and fairness. A report for all Canadian citizens*. Surveillance Studies Centre, Queen's University.

Lyon, D. (2022b). Surveillance. *Internet Policy Review*, *11*(4).

Lyon, D. (2022c). *Pandemic surveillance*. Polity Press.

Martin, K. (2016). Understanding privacy online: Development of a social contract approach to privacy. *Journal of Business Ethics*, *137*, *551–69*.

Mosco, V. (2014). *To the cloud: Big data in a turbulent world*. Routledge.

Nissenbaum, H. (2009). *Privacy in context: Technology, policy and the integrity of social life*. Stanford University Press.

OPC. (2021). Privacy and COVID-19 vaccine passports. (Joint statement of Canadian Privacy Commissioners). Office of the Privacy Commissioner of Canada. https://www.priv.gc.ca/en/opc-news/speeches-and-statements/2021/s-d_20210519/.

Pallitto, R.M. (2020). *Bargaining with the machine: Technology, surveillance and the social contract*. University Press of Kansas.

Parker, I., Montgomery, J., and Freeguard, G. (2021). "What place should COVID-19 vaccine passports have in society?," Ada Lovelace Institute. https://www.adalovelaceinstitute.org/report/covid-19-vaccine-passports/.

Phillips, J., and Mamuji, A. (2021, March 8). *Why is the uptake of digital contact tracing apps low? The Digital Global Health and Humanitarianism Lab has evidence-based answers and recommendations*. Dahdaleh Institute for Global Health Research, York University.

Puri, A. (2021). A theory of group privacy. *Cornell Journal of Law and Public Policy*, *30*(3), 477–538.

Rahwan, I. (2018). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology*, *20*(1), 5–14.

Renieris, E. (2021, April 5). *What's really at stake with vaccine passports*. Centre for International Governance Innovation. https://www.cigionline.org/articles/whats-really-stake-vaccine-passports.

Scassa T. (2021, March 30). *Interesting amendments to Ontario's health data and public sector privacy laws buried in omnibus bill*. www.teresascassa.ca.

Shils, E. (1956). *The torment of secrecy: The background and consequences of American security policies*. The Free Press.

Solano, J.L., Martin, A., Ohai, F., de Souza S., and Taylor, L. (2022). "Digital Disruption or Crisis Capitalism," Tilberg Institute for Law Technology and Society. Tilburg University. DOI: 10.26116/gdj-euaifund.

Srnicek, N. (2016). *Platform capitalism*. Polity Press.

Taylor, C. (2003). *Modern social imaginaries.* Durham, NC: Duke University Press.

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data and Society*, *4*(2).

Taylor, L., Martin, A., Sharma, G., and Jameson, S. (Eds.). (2020). *Data justice and COVID-19: Global perspectives*. Meatspace Press.

Toh, A., and Brown, D. (2020, June 4). How digital contact-tracing for COVID-19 could worsen digital inequality. *Human Rights Watch*. https://www.hrw.org/news/2020/06/04/how-digital-contact-tracing-covid-19-could-worsen-inequality.

Weller, T. (2012). The information state: An historical perspective on surveillance. In K. Ball, K. Haggerty, and D. Lyon (Eds.), *Routledge handbook of surveillance studies* (57–63). Routledge.

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, *30*(1), 75–89.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for human future at the new frontier of power*. Profile Books.

Zuboff, S. (2021, January 29). The coup we are not talking about: We can have democracy, or we can have a surveillance society, but we cannot have both. *New York Times*. https://www.nytimes.com/2021/01/29/opinion/sunday/facebook-surveillance-society-technology.html.

8

The Use of COVID-19 Exposure Notification Apps in Canada

A Deep Dive into Provincial Privacy Frameworks

Pierre-Luc Déziel

Introduction

In July 2020, the Government of Canada announced the launch of COVID Alert, a free and voluntary COVID-19 exposure notification application that allows users who tested positive for COVID-19 to notify people they recently came into contact with of their change in health status (Government of Canada [GC], 2020). By providing early warnings of possible exposure to the virus, the Government of Canada sought to encourage citizens to take the necessary precautions – such as self-isolation or getting tested – to limit the progression of the epidemic (GC, 2021). The application was made available to the provinces and territories for implementation within their jurisdiction and has been subsequently adopted by all of them except Alberta, British Columbia, Nunavut, and Yukon. COVID Alert was decommissioned in June 2022, after having sent no less than 456,359 notifications (GC, 2022).

The use of mobile applications in the fight against the spread of COVID-19 has raised several questions regarding their impacts not only on users' rights and freedoms but also on fundamental democratic values. Because they operate by using vast quantities of personal information, exposure notification apps have raised the specific issue of privacy (Bradford, Aboy, and Liddell, 2020). Many observers have thus expressed concerns that the collection of data that relate to the users' health status might lead to a new form of particularly intrusive governmental surveillance (Algorithm Watch, 2020; Amnesty International, 2020). Given that notification apps were new, many questions emerged as to exactly what type of data would be collected, who would be allowed to access and use those data, and what would happen with them once the epidemic was over. Moreover, the roles private actors, such as Google or Apple, played in the development of the applications remained unclear (Lyon, 2022). As David Lyon argues in this volume, the use of data-dependent technological "solutions" to the pandemic has nonetheless strengthened government-and-business partnerships and allowed platforms to enter new spheres, such as public health.

How governments, especially democratic ones, answered questions about the impact of COVID-19 apps on individual rights and freedoms is particularly important. For one thing, the distribution of personal information in a society has a tremendous impact on its internal power dynamics. Entities that have access to a significant amount of information about citizens can use this information to influence, manipulate, or even coerce them into acting in a way they otherwise would not. The protection of privacy is thus rightfully considered as a vector for freedom of expression, personal autonomy, and, consequently, as a necessary condition for democracy (Richard, 2021; Véliz, 2021).

It was therefore fitting that, when the Government of Canada launched COVID Alert, it also published several documents that provided detailed information about the application and explained how it would protect the users' right to privacy. From a privacy and data policy standpoint, the most important document was the privacy impact assessment (PIA), which provided background information about the application, listed the partners involved in its development, explained how it works, and provided the results of a thorough privacy analysis.

The PIA's main conclusion was that COVID Alert was designed to meet important public health goals while providing a high level of privacy protection. However, it is also important to note that the PIA clearly established that a component of COVID Alert, namely the one-time key distribution process, was to be administered at the provincial and territorial levels. The PIA repeatedly claimed that how information was to be handled at the provincial and territorial levels fell outside of the PIA's scope. In other words, the Government of Canada's privacy analysis of COVID Alert only covered how the information would be treated at the federal level and did not explicitly address how the provinces and territories would handle their side of the application.

If the Government of Canada acted transparently, the same cannot be said for the provinces and territories. Information on how they handled their side of COVID Alert is hard to come by. None of the provincial websites that recommended COVID Alert provide relevant and updated information about the provincial and territorial data governance policies. In fact, most websites merely redirected users toward the federal PIA. The lack of information on the provincial and territorial privacy frameworks therefore leaves citizens in the dark about the key elements of the application's overall privacy strategy.

The main contribution of the research expounded in this chapter is therefore to document and analyze the provinces' COVID Alert privacy frameworks to provide a complete assessment of the application's impact on its users' privacy. Information on the efforts deployed by the provinces and territories to assess the privacy implications of COVID Alert was collected through a series of freedom of information (FOI) requests submitted to the provincial health authorities.

The analysis of the information revealed by the FOI leads to two main conclusions. The first is that the provinces and territories do not seem to have developed independent and adequate privacy policies for dealing with the components of COVID Alert for which they were responsible. Most of the documents sent by the few provinces and territories that responded to the request were published by the Government of Canada and therefore deal with the federal side of the application. Moreover, many of the documents only provide general information about the application (e.g., PowerPoint slides or bullet point two-pagers) or focus on elements of COVID Alert that did not touch on privacy issues.

The second conclusion, which builds on the first, is that there seems to be a certain disconnect between how the privacy implications of COVID Alert are communicated to the public and what the few documents provided by the provinces and territories tell us about these implications. COVID Alert was often – if not always – presented by the provinces and territories as an application that did not involve specific privacy issues because it did not require the user to share any personal information. However, some documents reveal, for example, that there was an actual collection of personal information on the provincial and territorial side. The collection happened when authorities had to verify and confirm the identity of users to distribute one-time keys. To be sure, this collection of personal information did not raise serious privacy issues, but it is nevertheless puzzling to hear the provinces and territories repeatedly insist on the so-called fact that COVID Alert did not collect any personal information.

In this chapter, I begin by explaining how COVID Alert worked and show which components of the application were under the provinces' and territories' control and why they are important from a privacy standpoint. I then detail the data collection process used for this research and describe the type of documents that were provided by the province and territories in response to our FOI requests to show that the provinces and territories do not seem to have developed independent and adequate privacy strategies for dealing with the components of COVID Alert for which they were responsible.

I then further explain the above-mentioned disconnect between how the privacy implications of COVID Alert are communicated to its users and what the relevant documents provided by the provinces and territories tell us about these implications. Finally, I explore the possible causes and consequences of this disconnect and argue that it is mainly the result of a lack of adequate communication between the different governments involved in the implementation of COVID Alert, which led the provinces and territories to misunderstand their role and responsibilities in protecting the privacy of users. I also identify key lessons for governments to evaluate and communicate with citizens the privacy risks that the use of new technology may raise.

COVID Alert: The Importance of the Provinces' Privacy Frameworks

COVID Alert adopted a fully decentralized approach, which means that collected data was not stored in a central location but across several different locations. The application was built around three main elements:

the mobile app itself, the key server, and the one-time key distribution process. The Government of Canada was responsible for the first two elements, while the provincial and territorial governments oversaw the last one. Each element played an important role in the application and represented a key moment in the overall mechanism. The application had to be installed on a mobile device and relied on a specific type of exchange of information – which we sometimes referred to as "digital handshakes" – between devices that are in proximity of one another.

When COVID Alert was activated, the device would send out rolling proximity identifiers (RPIs) to other devices while simultaneously listening for RPIs from other devices, which would be stored directly on the user's device. RPIs were random codes generated every five to twenty minutes by a random temporary exposure key (TEK), which was generated every day by the application. Because they were by design meant to be public and shared, the RFIs were not identifiable and did not contain any personal information. Moreover, the fact that the RFIs and the TEK were not static but created and refreshed on a regular basis significantly diminished the risk of identification of the users (GC, 2021).

The notification process had to be initiated by a user who received a positive diagnosis of COVID-19. The user asked the province or territory where they lived for a one-time key (OTK). The province or territory generated the OTK via a portal linked to the federal key server and distributed it to the user. The key was then entered by the user on their mobile phone. When the key was validated, COVID Alert asked if the user consented to send their TEKs to the key server. Once they reached the key server, the TEKs became "diagnosis keys," which could then be downloaded by other users to start the notification process. The notification process worked like this: on a regular basis, the app automatically downloaded new diagnosis keys available on the key server, regenerated the RPIs associated with them, and compared them to the RPIs stored on the users' device. If there was a match between the RPIs downloaded via the diagnosis keys and the RPIs stored on the device, it meant that the user's device could have been in the proximity of a device whose user had tested positive for COVID-19.

One of the key features COVID Alert relied on was the separation of responsibilities between the federal government and the provincial and territorial governments: the Government of Canada was responsible for developing the application and managing the key server, while the provinces and territories were responsible for distributing the OTKs to users who tested positive for COVID-19. The Government of Canada was confident that COVID Alert provided a considerable level of privacy protection to its users. However, it also made it clear that those conclusions only applied to the elements of the application that fell under its responsibility.

It is important to note that the PIA clearly stated that it focused exclusively on the privacy impact of the application that related to the information "under the control of the Government of Canada, including the data that will be transmitted to the Government of Canada Server (key server)" (GC, 2021, "Scope"). It also explicitly specified that it did not "examine any information under the control of the Provinces/Territories (PTs), such as the mechanism through which one-time keys are distributed to individuals for PTs who distribute one-time keys themselves" (GC, 2021, "Scope"). Moreover, if the federal side of the application provided the OTKs to the provinces or territories, the decentralized structure made it impossible for the Government of Canada to know "who that key is associated with (in other words who the person is who will receive this key)" (GC, 2021, "How are individuals notified of exposure to someone who tested positive"). In other words, the Government of Canada only handled keys – which are basically alphanumeric codes – while the provinces were responsible for matching a given individual who tested positive with a specific OTK.

This is why the Government of Canada expected that "each province will protect the information in accordance with the requirements imposed on them by applicable provincial or territorial legislation" (GC, 2021, "Is there a collection of personal information/data" and "One-Time Keys and API Tokens"). Therefore, if the PIA concluded that COVID Alert provided a high level of privacy to the users, it is also clear that this conclusion only applied to the federal side of the application and that it did not cover the elements under the provinces' responsibility. Since it is the provinces and territories that were handling the most sensitive information by pairing individuals, their official COVID-19 diagnosis, and the OTK, it is also evident that how the provinces protected this information is critical for understanding the privacy impact of COVID Alert as a whole.

Nothing to See: A Deep Dive into the Provincial Data Privacy Frameworks

To gather information about how the provinces and territories protected the information under their control, we submitted a series of FOI requests to the Departments of Health of all the provinces and territories that adopted the application, except Manitoba.[1] In certain cases, such as Prince Edward Island, the Access and Privacy Office recommended that we submit additional requests to other departments likely to have the relevant documents, such as the Department of Finance. In this particular case, we submitted supplemental requests. In other cases, such as in Ontario and Newfoundland and Labrador, time constraints, unfortunately, did not allow for this type of follow-up.

The FOI requests that we submitted had two broad sections. The first provided context for the request and explained our view that certain components of COVID Alert were under the control of the provinces. The second listed the requested documents. The list was long and, to a certain extent, demanding. Our main objective was to get as much information as possible by requesting documents that could touch any privacy or data governance aspect of COVID Alert. The FOI requests were submitted to the provinces in February and March 2022 via the appropriate channels. Here is the list of the documents that we asked for:

- The data management and governance policies for any type of data and personal information related to the provincial component of the COVID Alert;
- Documents detailing the mechanisms for generating and assigning one-time identification keys and explaining how the keys are – or are not – matched to positive COVID diagnoses within the province;
- The policies related to the management of access to this information and the security measures (technical, organizational, and physical) applied to the information;

- The memoranda of understanding (MOU) entered into with the Government of Canada regarding the protection of single-use keys;
- The nature of the personal information or data collected, the purposes relating to their collection and use, and the policies related to the communication of this information;
- All expert reports and all opinions, legal, technical, or otherwise, related to the privacy impact assessment in connection with the administration, implementation, and management of the provincial component of the COVID Alert application;
- All documents relating to the storage (formats, durations, and purposes) of data and personal information under the control of the province and relating to the COVID Alert application; and
- The list of actors, companies, and organizations, public or private, involved in the management of data and information related to the main component of the COVID Alert application.

The results of the FOI were not consistent across the provinces and territories. Some provinces and territories did not return any documents, while others provided large and voluminous files.[2] In the latter case, most of the documents were produced by the Government of Canada and did not provide any meaningful insights into how the provinces and territories addressed the privacy issues that could stem from the management of their side of COVID Alert. Moreover, many documents did not directly focus on privacy or only provided information that was already widely available online.

Three provinces (Saskatchewan, Ontario, and New Brunswick) did not return any documents as a result of the FOI. Nova Scotia, Prince Edward Island, Quebec, and the Northwest Territories provided meaningful answers to the FOI. However, as stated earlier, the documents provided were often not helpful in understanding if and how the provinces built their privacy frameworks for COVID Alert. Table 8.1 lists and briefly describes the different documents these provinces and territories sent.

As we can see, despite the large volume of files submitted by the provinces, most of the information shared did not specifically address privacy issues, nor the role of the province in managing COVID Alert.

Table 8.1 Type of document sent by Canadian provinces

	Type of document submitted and description
Nova Scotia	• Dashboard for COVID Alert's Call Centre and its Generic Email Weekly reports; • PowerPoint presentation by Health Canada explaining how COVID Alert works; • Redacted emails that follow Health Canada's presentation and by which the files on COVID Alert's Call Centre and Generic Emails Service were sent to the province; • Two-page template used by the provinces to identify the provincial representatives to whom meeting invitations should be sent by Health Canada; • A document entitled "COVID Exposure Notification Service Concept" written by the Treasury Board of Canada; • Canada–Nova Scotia COVID-19 Exposure Notification System Agreement / MOU; • Two-page document that establishes the terms of reference for a "Federal, Provincial and Territorial Public Health Working Group on the National COVID-19 Exposure Notification App.
Prince Edward Island	• One-page, low-quality scan of a "Service Map" produced by the Canadian Digital Service that illustrates the role and responsibilities of different actors in managing the application; • Privacy Checklist for COVID-19 Initiatives produced by Health Canada; • Canada–Prince Edward Island COVID-19 Exposure Notification System Agreement / MOU; • PowerPoint presentation by Health Canada and Public Health Agency of Canada detailing the marketing plans for COVID Alert; • PowerPoint presentation by the Canadian Digital Service providing general information about the application's rollout strategy; • A 124 page document that presents the results of COVID Alert's use, such as the number of downloads, IOS users, and Android users; • Canada-NT COVID-19 Exposure Notification System Agreement / MOU; • Privacy Impact Assessment conducted by the Health Privacy Unit of the Department of Health and Social Services of the Government of the Northwest Territories.
Quebec	• Two-page briefing note on COVID applications operating with geolocation at the international level; • A short, bullet point summary of COVID Alert; • Canada-Québec COVID-19 Exposure Notification System Agreement / MOU; • A point-by-point response to the list of documents we requested through the FOI.

There are, however, exceptions. Three documents are, for our purposes, relevant.

The first is the MOU, which is explicitly required by the PIA and takes the more specific name of "COVID-19 Exposure Notification System Agreement." All the provinces and territories that replied to our FOI requests submitted the MOU they signed with the Government of Canada. As we will see later in this chapter, the MOU is highly relevant for our purposes because it provides interesting information about the duties and responsibilities of the provinces with respect to the protection of users' privacy.

The second relevant document is a PIA conducted by the Health Privacy Unit of the Department of Health and Social Services of the Government of the Northwest Territories. This PIA repeatedly insists that the application did not collect any personal information and that users were not required to provide any identifying information to use the application. The document also states that staff in the Department of Health and Social Services would "generate one-time keys for issuance to individuals to facilitate exposure notification" and that the OTKs would be "provided to the individual for entering into the COVID Alert application, prior to use." However, the PIA does not explain how the OTKs would be distributed to the user, nor how the identity and health status of the user would be verified. As we will see later, this point may prove to be important because the MOU clearly states that the provinces and territories have the responsibility to distribute OTKs only to users who tested positive for COVID-19. How the identity and the health status of the users will be verified and confirmed is not explained in the PIA.

The last relevant document was sent by the Government of Quebec. It addresses, point by point, the list of documents we requested through the FOI. The response to all but one point is that there are no relevant privacy documents to release because the province simply did not retain any personal information. The only point for which a more substantial answer is provided is the one on the mechanisms used for generating and assigning OTKs. Here, the document explains that users can voluntarily call the telephone number provided by the application to get their OTK. The operator will confirm the user's identity and verify that they tested positive for COVID-19. The operator will generate the key and communicate it to the user. The document stresses that at no point will the operator record, retain, or keep any personal information on the user. I will get back to this point in the next section, when I argue that the fact that the province does not retain personal information technically implies that it had to at least collect it in the first place.

Despite these few exceptions, what the documents sent by the provinces and territories demonstrate is that they have not elaborated and implemented effective privacy policies for the components of the application under their responsibility. It thus seems that the provinces and territories have taken for granted the claim, widely advertised by the Government of Canada, that COVID Alert did not collect any personal information and therefore did not undertake a serious investigation of the privacy issues that distributing OTKs to individuals who tested positive for COVID-19 may raise. In fact, only Quebec and the Northwest Territories have addressed this question, even if briefly. I return to this issue in the next section.

The Disconnect between the Message and the Reality of COVID Alert

In this section, I explain why there are good reasons to think that the management of the provincial side of COVID Alert required the collection of personal information from the users. This finding is significant because it contradicts the widely advertised notion that COVID Alert did not collect any personal information. Before going any further, a quick precautionary comment is necessary. I recognize that my conclusion on this point rests only on a few documents provided by the provinces and, as a consequence, is not supported by a large number of observations drawn from the result of the FOI. However, it nevertheless seems clear, at least from a purely deductive point of view, that some form of collection of personal health information would occur at the provincial level. As we will see, the final evaluation provided by Health Canada on COVID Alert also raises several points that tend to support the findings.

In my opinion, the collection of personal information that must have taken place at the provincial and territorial level would have occurred when health authorities generated OTKs and assigned them to users who tested positive for COVID-19. As we have seen above, users had to enter the OTK in the application to start the notification process. The distribution and assignment of the OTKs were elements of COVID Alert under the control of the provinces and territories. The MOUs provided by Quebec, Nova Scotia, the Northwest Territories, and Prince Edward Island offer additional information as to what is expected from the provinces and territories with respect to assigning the OTKs to users and therefore paint a clearer picture of what must have happened when they administered their side of the application.

Section 5 of the MOU clearly explains that the provinces and territories had the responsibility to distribute OTKs to users who tested positive for COVID-19. In doing so, provinces had to "ensure … that queries to request the OTKs are not submitted unnecessarily or maliciously and the OTKs are only distributed to individuals in [the province or territories] who test positive for COVID-19 and that only one OTK will be requested of each positive test result of an individual." In other words, provinces and territories were required to confirm both the identity of the users who request OTKs and that they tested positive for COVID-19. The goal was to prevent users who did not test positive for COVID-19 from spreading false information to other COVID Alert users by initiating useless notification processes.

The authentication of users and the verification of their official diagnosis must respect the measures described in Section 7, "Due Diligence," of the MOU. Here, the MOU clearly points out that it was expected that the provinces and territories develop and implement security and privacy measures to protect the users. It explicitly states that provinces and territories had to "employ reasonable and necessary physical, administrative and technical safeguarding measures in order to abide by all legislation and policies applicable to them, along with any applicable industry standards or best practices related to privacy and security in the deployment of similar systems." It is thus clear that protecting the users' privacy was not a responsibility that fell exclusively on the federal government; provinces and territories also had to adopt, by themselves, measures to protect the users.

This conclusion stems from the fact that, to authenticate users and verify their health status, provinces and territories necessarily had to collect *some* information about the users; it is hard to imagine how the provinces could fulfill these duties without collecting at least some information that confirms the identity of the individual and their health status. On this point, it might be important to note that personal information is defined by the Privacy Act as "information about an identifiable individual," which includes basic, nonsensitive demographical information that allows for the identification of the person. Most importantly, provinces and territories all have laws that adopt an almost identical definition of personal information. Therefore, any information used to confirm the identity of COVID Alert users was, by definition, personal information.

Moreover, Section 5 ("Roles and Responsibilities") of the MOU directly alludes to the fact that information about the identity of the users may be collected when it explicitly states that the provinces and territories "will not retain the OTKs or information linking the identity of the individuals who tested positive to the OTKs." On the same point, let's recall Quebec's answer to our FOI, which stated that the province did not produce any relevant privacy documentation or analysis because it "does not retain any personal information" on the users.[3] The mere fact that personal information – that is, information that links the identity of the individual to other pieces of information – should not be retained strongly suggests that personal information has been collected by the provinces.

Trying to Make Sense of the Provinces' Inaction: Learning from the COVID Alert Experience

Given what was said above, it is easy to see why claims to the effect that provinces and territories did not have to produce any written privacy policies because the COVID Alert did not retain any personal information are at best disappointing.[4] For reasons I will discuss here, the collection of personal information by the province does not appear to be, in and of itself, particularly invasive or problematic from a legal standpoint. What is interesting and problematic is not the collection of personal information per se but this insistent notion that COVID Alert did not collect any personal information when it clearly did.

What explains this disconnect? The most plausible scenario is that provinces and territories did not fully understand the extent to which administrating their side of COVID Alert would require the collection of personal information. It is reasonable to suggest that, given the novelty and the complexity of the technology, the provinces and territories have taken for granted the Government of Canada's claim that COVID Alert did not collect personal information. Such deference towards the federal government is not entirely surprising or unreasonable, especially when one considers that provinces and territories were overwhelmed by the pandemic.

It is also interesting to note that, in its final report on COVID Alert, Health Canada concludes that the provinces have not been adequately consulted during the development of the application and that the communication between the provinces and the Government of Canada was essentially "a one-way exchange of information from the federal government" (Office of Audit and Evaluation Health Canada, 2022, 21). Health Canada also claims that time constraints have seriously hampered the federal government's ability to assess the provinces' needs. The application was therefore introduced "without a full understanding of the needs of individual [provinces] and each jurisdiction's capacity to fulfill their needed role (i.e., giving out OTKs to people who tested positive for COVID-19)" (21). These findings not only support my claim that assigning OTKs was a key responsibility of the provinces and territories that may have involved the collection of personal information, but it also suggests that the provinces did not fully understand its implication. In other words, it seems like the disconnect between what was communicated to the public about COVID Alert and what happened is the result of a lack of understanding rather than a blatant lie.

As we said earlier, the collection of personal information by the provinces does not appear to be, in and of itself, particularly problematic. The authentication of citizens for the delivery of government service is a routine task that can readily be tailored to meet basic privacy principles. There is no need to collect particularly sensitive information, and confirming the identity of users to prevent false or malicious notification processes is a reasonable and acceptable purpose. Moreover, the collection of personal information is voluntary, thereby respecting individual consent.

But that this collection of information does not seem to raise particularly daunting privacy issues does not mean that these issues should not be taken seriously. For one, even if the collection is limited to basic identification information, it should be kept in mind that personal information was collected to be linked with personal *health* information, which normally attracts a higher level of privacy considerations and protection. Second, basic information ought to be taken seriously because even the most trivial information can be used, by itself or with other pieces of data, to generate new insights into a person's life or personality (Kosinski, Stillwell, and Graepel, 2013). Third, privacy breaches can occur while conducting more basic tasks, such as faxing information to a patient (Nova Scotia Health Authority and Private Practice Physicians (Re), 2016), or through the mundane incidents of everyday life (Public Hospital (Re), 2017), such as misplacing a dossier (Department of Health (Re), 2020) or losing a mobile device (British Columbia Health Authority Privacy Breach Management (Re), 2015). Privacy breaches are therefore prevented not only by the use of sophisticated encryption algorithms but also by the careful consideration of the importance of administrative measures such as the adequate training of health professionals and the elaboration of comprehensive privacy policies (Review of the Electronic Health Information System at Vancouver Coastal Health Authority Known as the Primary Access Regional Information System, 2010; Northwest Territories Health & Social Services Authority – Yellowknife Region (Re), 2020).

Provinces and territories – and by extension the Government of Canada – have failed to provide even such basic privacy protection to their citizens. The fact that they neglected to elaborate cohesive privacy policies for the personal information they did collect suggests that provinces did not take seriously the privacy implication of COVID Alert. And this is precisely the issue. As I said, the collection of personal information is not, by itself, particularly problematic; it is the fact that it was not recognized as such that is the problem. Because the documents produced by the provinces and territories for the FOI requests did not provide any significant information on the actions taken by the provinces to protect the privacy of COVID Alert's users, it is impossible to know exactly who proceeded to the collection of personal information, what type of information was collected, what was the stated purpose, or how – and if – the collected information was deleted or destroyed. Moreover, it is impossible to know the physical, administrative, or technical measures taken by the provinces to protect the privacy of the users, as required by the MOU. In other words, we do not have any relevant information as to how the provincial and territorial governments handled the personal information under their control while administrating COVID Alert.

Given that 456,359 notifications were sent through COVID Alert, and therefore that an equal number of OTKs were distributed, it is impossible for the provinces and territories to account for how they managed the collection of hundreds of thousands of pieces of personal information. This is unacceptable for a democratic government. But the fact that provinces did collect personal information while the Government of Canada advertised that COVID Alert did not collect any personal information raises issues that go beyond accountability.

Luckily for the Government of Canada, and the provinces and territories, there has been no security incident with COVID Alert – at least none that was made public. However, it is not hard to imagine that, in the event of an actual incident, the lack of accurate and complete information for users could be perceived as dishonest, malevolent, or incompetent. Transparency is not only a fundamental privacy principle; it is also necessary to the maintenance of public trust during a health crisis. Communication that is not transparent, truthful, or accurate may undermine trust in the government, nourish public anxiety, and lead citizens to look for alternative sources of information (Hyland-Wood et al., 2021; Prah Ruger, 2020; Siegrist, De Campos-Rudinsky, and Unurrage, 2021; Siegrist and Zingg, 2014). Because trust and transparency were presented as central components of the Government of Canada's strategy to boost the use of COVID Alert and contribute to its effectiveness, the disconnect between how the application was presented to the users and how it operated could have been particularly problematic.

These findings raise the interesting question of how privacy risks can and should be communicated, especially when they involve the use of a new technology by governments. What is clear in the case of COVID Alert is that both the Government of Canada and the governments of the provinces and territories have failed to provide exact information about how the application worked and the privacy risks associated with its use. Advertising the idea that COVID Alert did not collect *any* personal information was false and any incident that would have exposed this fact could have led to particularly disastrous consequences.

What lessons, then, should we draw from the COVID Alert experience? What should we learn and how can we best prepare for the next time governments have to advertise new technologies to their citizens?

In my opinion, the main lesson is that governments – and especially democratic ones – should never take for granted that an application does not collect any personal information without verifying, for themselves, how the technology works. Governments and public organisms are by law responsible for the technologies they use and the personal information under their control; they should not blindly accept the results of a PIA they did not conduct, even when they operate under stress and with significant time constraints. Advertising that a technology does not involve any privacy risks is too strong a claim to make without careful examination and verification. With COVID Alert, it seems that the provinces and territories simply took for granted the Government of Canada's claim that the application did not collect any personal information and did not conduct an independent analysis of the technology as a whole and what their responsibilities would require in terms of data collection.

The second lesson, which draws on the first, is that when a technology requires the collaboration of multiple entities that have different duties and responsibilities, a single PIA that considers *all* components of the application should also be conducted. Governments should refrain from partitioning a PIA into multiple documents because the conclusion of a single PIA may not apply to the technology as a whole and may thus provide an inaccurate picture of the privacy issues involved. Given our first lesson, there are two possible ways to operate. Either a single PIA is conducted with the participation of all parties involved or each party conducts its own PIA that takes into account the application as a whole. In the latter case, results should be compared to make sure there are no contradictions or misunderstandings. With COVID Alert, part of the problem stems from the fact that the PIA was conducted by the Government of Canada and it only focused on the elements of the application under its control; even if its analysis did not apply to COVID Alert as a whole, its conclusions were nevertheless taken as a valid for all components and parties involved.

Finally, it seems that governments should be particularly prudent when they provide information about technologies for which the participation of the public is a necessary condition of its effectiveness. Applications like COVID Alert only work if a significant number of people adopt it (Abueg et al., 2021; Braithwaite et al., 2020; Vogt et al., 2022). It might be tempting, then, to oversell the application by embellishing the privacy protection it offers to users. Users who do not trust that their information will remain confidential or worry about possible misuse of the information collected by the application could then refrain from adopting it (Altman et al., 2020). Alleviating the privacy concerns of potential users is therefore fundamental to boosting adoption and, by extension, effectiveness. However, it must also be kept in mind that providing meaningful, accurate, and truthful information on how the application works builds trust in the government and, ultimately, contributes to the effectiveness of the application (Lavorgna et al., 2021; Ranisch et al., 2021; Seto, Challa, and Ware, 2021).

Conclusion

My main objective in this chapter was to provide a complete assessment of COVID Alert's impact on its users' privacy. Through a series of FOI requests, I was able to gain more information on how the provinces have protected users' privacy while managing the components of the application that fell under their control. The documents submitted by the provinces lead to two conclusions. First, the provinces did not elaborate meaningful privacy frameworks to protect the privacy the users. Second, there are good reasons to believe that COVID Alert did collect personal information. This finding is significant because it contradicts the widespread claim of the Government of Canada and the provinces and the territories that the use of COVID Alert does not raise any privacy issues because it does not collect any personal information. The collection of personal information would have taken place when the provincial health authorities confirmed the identity of the users who requested a OTK and verified that they tested positive for COVID-19.

I argued here that the disconnect between how COVID Alert was advertised and how it actually worked results mainly from a lack of effective communication between the federal government and the provinces and territories. As a consequence, provinces and territories did not have a clear understanding of what their roles and responsibilities were in respect to COVID Alert. Using COVID Alert as a case study for how privacy risks should be evaluated and communicated by governments who roll out new technologies in emergency contexts, I identified three main lessons. First, governments should always conduct their own independent privacy analysis and refrain from taking for granted conclusions that other partners involved might provide. Second, PIAs should consider an application as a whole and should not focus only on a selected number of its component. Finally, governments and public organisms should be particularly prudent when they communicate with the public and take all necessary precautions to be transparent and provide their citizens with accurate information about how the technology works.

Notes

[1] We excluded Manitoba on the basis that it was the only province and territory where the ATI request could not be submitted online. Contrary to all the other departments in Manitoba, the Department of Health does not provide an online form to request information and does not have an online portal. The person in charge of access to information and privacy is the only one to not even provide an email address. See https://www.gov.mb.ca/fippa/wheretosend/index.html.

[2] All the documents are on file with the author and available here: https://drive.google.com/drive/folders/1IDTOacV6RcIb2T_Q90LOEyPCZEUS3nfr?usp=sharing.

[3] In French, the answer is “Aucune information personnelle n’est conservée.”

[4] The document explains, in French, that “Il n’y a pas de politique écrite, considérant que la règle est de ne conserver aucune donnée personnelle des appelants.”

References

Abueg, M., Hinch, R., Wu, N., Liu, L., Probert, W., Wu, A., Eastham, P., Shafi, Y., Rosencrantz, M., Dikovsky, M., Cheng, Z., Nurtay, A., Abeler-Dörner, L., Bonsall, D., McConnell, M.V., O’Banion, S., and Fraser, C. (2021). Modeling the effect of exposure notification and non-pharmaceutical interventions on COVID-19 transmission in Washington state. *NPJ digital medicine*, *4*(1), article no. 49.

Algorithm Watch. (2020). *Automated decision-making systems and the fight against COVID-19 – Our position*.

Altmann, S., Milsom, L., Zillessen, H., Blasone, R., Gerdon, F., Bach, R., Kreuter, F., Nosenzo, D., Toussaert, S., and Abeler, J. (2020). Acceptability of app-based contact tracing for COVID-19: Cross-country survey study. *JMIR mHealth and uHealth*, *8*(8), e19857.

Amnesty International. (2020). *Joint statement: States use of digital surveillance technologies to fight pandemic must respect human rights.*

Bradford, L., Aboy, M., and Liddell, K. (2020). COVID-19 contact tracing apps: A stress test for privacy, the GDPR, and data protection regimes. *Journal of Law and the Biosciences*, *7*(1), lsaa034.

Braithwaite, I., Callender, T., Bullock, M., and Aldridge, R.W. (2020). Automated and partly automated contact tracing: A systematic review to inform the control of COVID-19. *The Lancet Digital Health*, *2*(11), e607-e621.

British Columbia Health Authority Privacy Breach Management (Re). BCIPC 66. 2015.

de Campos-Rudinsky, T.C., and Undurraga, E. (2021). Public health decisions in the COVID-19 pandemic require more than "follow the science." *Journal of Medical Ethics*, *47*(5), 296–99.

Department of Health (Re). NUIPC 14. 2020.

Government of Canada. (2020). *New mobile app to help notify Canadians of potential COVID-19 exposure now available* [Press release]. https://pm.gc.ca/en/news/news-releases/2020/07/31/new-mobile-app-help-notify-canadians-potential-covid-19-exposure-now.

Government of Canada. (2021). *COVID Alert: COVID-19 exposure notification application privacy assessment*. https://www.canada.ca/en/public-health/services/diseases/coronavirus-disease-covid-19/covid-alert/privacy-policy/assessment.html.

Government of Canada. (2022). *COVID Alert performance metrics.* https://www.canada.ca/en/public-health/services/diseases/coronavirus-disease-covid-19/covid-alert/performance-metrics.html.

Hyland-Wood, B., Gardner, J., Leask, J., and Ecker, U.K. (2021). Toward effective government communication strategies in the era of COVID-19. *Humanities and Social Sciences Communications*, *8*(1).

Kosinski, M., Stillwell, D., and Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*(15), 5802–805.

Lavorgna, A., Rekha, G.S., Ugwudike, P., Carr, L., and Benitez, Y.S. (2021). To app or not to app?: Understanding public resistance to COVID-19 digital contact tracing and its criminological relevance. *Law, Technology and Humans*, *3*(2), 28–45.

Lyon, D. (2022). *Pandemic surveillance*. Polity Press.

Northwest Territories Health and Social Services Authority – Yellowknife Region (Re). NTIPC 37. 2020.

Nova Scotia Health Authority and Private Practice Physicians (Re). NSOIPC 16. 2016.

Office of Audit and Evaluation Health Canada and the Public Health Agency of Canada. (2022). Evaluation of the national COVID-19 exposure notification app.

Prah Ruger, J. (2020). Positive public health ethics: Toward flourishing and resilient communities and individuals. *The American Journal of Bioethics*, *20*(7), 44–54.

Public Hospital (Re). OIPC No. 259. 2017.

Ranisch, R., Nijsingh, N., Ballantyne, A., van Bergen, A., Buyx, A., Friedrich, O., Hendl, T., Marckmann, G., Munthe, C., and Wild, V. (2021). Digital contact tracing and exposure notification: Ethical guidance for trustworthy pandemic management. *Ethics and Information Technology*, *23*, 285–94.

Review of the Electronic Health Information System at Vancouver Coastal Health Authority Known as the Primary Access Regional Information System. BCIPC 530. 2010.

Richards, N. (2021). *Why privacy matters*. Oxford University Press.

Seto, E., Challa, P., and Ware, P. (2021). Adoption of COVID-19 contact tracing apps: A balance between privacy and effectiveness. *Journal of Medical Internet Research*, *23*(3), e25726.

Siegrist, M., and Zingg, A. (2014). The role of public trust during pandemics. Implications for crisis communication. *European Psychologist*, *19*(1), 23.

Véliz, C. (2021). *Privacy is power: Why and how you should take back control of your data*. Bantam Press.

Vogt, F., Haire, B., Selvey, L., Katelaris, A.L., and Kaldor, J. (2022). Effectiveness evaluation of digital contact tracing for COVID-19 in New South Wales, Australia. *The Lancet Public Health*, 7(3), e2.

Conclusion
Democracy as an Artifact

Julia Rone and Cecilia Biancalana

Artificial Democracy: What Have We Learned?

Data are all around us and are part of our everyday life. When we use e-commerce sites, communicate with our friends and relatives (and even with the public administration), or simply walk through a "smart" city, we leave behind us the traces that constitute our "data double": a version of ourselves that more and more rules our lives. Our physical existence can no longer be separated from our online existence, as the two, through data, compenetrate each other. Politics, and in particular democratic politics, is not exempt from these processes.

This book gives a novel and original perspective to the multiple ways big data collection, analytics, and political uses have transformed contemporary democracies and redrawn the relations between citizens, political parties, governments, and private corporations. Most prominent discussions on big data's democratic impact so far have tended to focus on scandals such as the 2013 Edward Snowden revelations on governmental mass surveillance online (Bauman et al., 2014) or Cambridge Analytica and the few corporate actors that have acted in ways that are damaging for democracy within a broader context of "surveillance capitalism" (Dowling, 2022; Manokha, 2018; Zuboff, 2019). In addition, extensive research has outlined the importance of big data and microtargeting for campaigning by political parties (Anstead, 2017; Stromer-Galey, 2019; Tufekci, 2014; Zuiderveen Borgesius et al., 2018). What we have missed so far, however, has been an attempt to offer a broader analysis of how these different types of actors relate to each other in their uses of big data within a broader data ecosystem that is fractured by differences in national legislations, political systems, and cultures (Kefford et al., 2022).

The contribution of the current book is precisely such a comprehensive approach to big data's effect on democracy, outlining more specifically the ways big data has transformed politics (election campaigns and inter-party interactions), policies (with a specific focus on government policies with regard to health data during the COVID-19 pandemic), and finally, democratic polities themselves (including the epistemic foundations of democracy and the very ways we conceptualize the relationship between individuals, society, and government).

Rather than focusing on a few bad players, the different contributions of the book outline the very conditions for the functioning of surveillance capitalism, including the current legal regimes that treat data above all as personal data while failing to consider data's collective aspects (see Trudel's contribution to this volume); the reluctance of political parties in Canada to regulate the use of personal data and microtargeting by political parties (see Bennett's contribution to this volume); as well as the general inadequacy of foundational notions such as "the social contract" to describe relations between governments and citizens that are increasingly mediated by private corporations (see Lyon's contribution to this volume). These are all fundamental problems that define the level playing field and lie at the root of many of the problematic data practices we observe, including the increasing prevalence of closed party communication targeting individuals rather than other forms of more "open communications" that facilitate citizens' ability to make informed and, at least to some extent, impartial choices in a democracy (see Blais's contribution to this volume).

A second major contribution of this book is the way it has managed to move attention away from US and UK context and to draw attention to developments in other democratic countries such as Canada and Italy, as well as often overlooked levels of governance such as the local and regional. Several of the contributions to the book have a strong empirical focus and show how many of the trends and phenomena observed in relation to big data use in the US and UK manifest in radically different ways in other contexts.

For example, Biancalana's original ethnographic study of electoral campaigning at the local level in Turin, Italy, shows how practices of canvassing and microtargeting well known from the US context fail to translate to the Italian context, where people are highly wary of opening their doors to strangers and older, more traditional forms of campaigning still hold greater appeal. This kind of empirical research also raises important questions about parties' real-life use of big data, as well as about the novelty and success of new forms of political campaigning, which sometimes acquire mythical dimensions in academic research (Baldwin-Philippi, 2017; Simon, 2019) while remaining underused in practice.

The importance of studying the local and regional level also manifests in Déziel's study of provincial data governance structures in Canada in relation to COVID-19 exposure notification apps, which puts into context and problematizes calls for more state sovereignty over data governance (as seen in Trudel's contribution to this volume, for example). How do states' prerogatives and ambitions to achieve digital sovereignty over data, for example, play out in the context of federal governments and power devolution? While explored in the context of Canada, this question about scaling "data sovereignty" has wider implications when it comes to countries such as Germany or even transnational bodies such as the European Union.

Finally, a key strength of the book is how it addresses this rich variety of questions from an interdisciplinary perspective, bringing together insights from computer science, sociology, political theory, political ethnography, political communication, and law. And indeed, it is only by combining these three approaches that the authors can come to important insights about not only the current transformations of politics, policy, and polities but also about the desirable changes in these trends, the potential normative directions to be taken in each of the contexts explored, and the research opportunities for social sciences scholars (see Ouellet and Dufresne's contribution to this volume).

In what follows, rather than looking at individual contributions, we discuss some cross-cutting key issues raised by the book, such as the importance of transparency, accountability, and sovereignty; the need for ongoing political research on evolving and emerging political data practices; and the contrasts and similarities between the development of, and democratic concerns relating to, applications for electoral micro-targeting (relatively institutionalized, although variable by context and still evolving) and for public health in the context of the pandemic (experimental, crisis-driven, probable benchmark).

We then move on to discuss topics less discussed in the book, outlining avenues for future research on the basis of current contributions. We will draw attention, more specifically, to the role of private actors such as digital platforms and AI companies in developing practices and policies, as well as to the role of civil society actors such as citizen movements or NGOs, but also news media in critiquing but also using data as part of a "counterpower" strategy, adding additional key actors to the ones discussed in the book.

Data Democracy: Key Themes and Topics

Remedying Big Data's Dangers to Democracy: Why Data Protection Is Not Enough

A key topic that runs throughout several of the contributions in this book is the importance of the legal regimes regulating the collection, analytics, and use of data for public health protection and political campaigning alike. These legal regimes are presented not simply as an objective given but as highly contested and political. To begin with, Bennett's analysis of political party cartels preventing the passing of data protection legislation reveals clearly the politics of law making when it comes to data governance. But crucially, as Trudel shows, the very epistemic foundations behind current attempts to regulate political campaigning and health data also have strong political implications. Treating data as a personal, rather than collective, resource fits coherently within a liberal political tradition that focuses on individual rights and places the onus on the individual to defend these rights. The multiple inequalities of data production and extraction, and the ways marginalized communities are particularly vulnerable to data extraction, thus remain overlooked (Cheeseman, 2022). Thinking about data protection as concentrated on protecting privacy as an individual right similarly overlooks the broader consequences for democracy of aggregating and analyzing individual data (Zuboff, 2019), as well as the ontologically collective nature of data as a resource that reveals as much about individuals as about the social relations between them (see Trudel, this volume; also see Solove, 2013).

Thus, if we want to change the relationship between data and democracy, it is not enough to improve data protection. What is needed is a more fundamental rethinking of how we conceive data to begin with. The European General Data Protection Regulation (GDPR), often presented as a gold standard, relies on several cornerstones such as consent on the part of the individual but also transparency on the part of data operators about why they are collecting data and for what purposes the data will be used, as well as accountability – the idea that data operators could be held accountable in the case of a breach. Each of these current cornerstones of data protection have been strongly criticized for more than a decade now.

To begin with, to enforce privacy rights through consent, it is, of course, important to know what uses personal data is being put to; therefore, transparency and accountability have been considered key aspects of data protection. In her analysis of recommendations made by the official bodies investigating the Brexit referendum process, Katharine Dommett (2020) noted that they consistently demanded more transparency in political campaigning. Still, the official commissions rarely specified what type of transparency is expected, with each having diverse demands for financial, source, data, or targeting transparency. What is more, the form and specifics of transparency – who the targeted audience is; how discoverable, comprehensible, and reliable the information is – was also rarely specified. This meant that general calls for transparency often left it to private corporations and political parties alike to decide what to make transparent, when, and how (Dommett, 2020). And as seen in Bennett's contribution in this volume, not only private corporations, but also political parties are reluctant to voluntarily disclose detailed information about their operations, not to mention to subject themselves to tight privacy legislation.

But what is to be done about the current focus on data protection as the solution to some of the concerns big data poses to democracy? In a sense, the fact that Canada has not yet managed to introduce laws on political campaigning might open up opportunities to rethink forms of data governance at a deeper level, taking into account ideas of data as a collective resource that states have a sovereign right to manage (as suggested by Trudel). At the same time, as tempting as it might be, opposing data sovereignty approaches to the current dominant paradigm of individualized data protection also comes with substantial pitfalls and caveats.

Problematizing Digital Sovereignty

Digital sovereignty has been most broadly defined as a "form of legitimate, controlling authority over – in the digital context – data, software, standards, services, and other digital infrastructure" (Floridi, 2020, 2). While being in many respects an "empty signifier" (Lambach and Oppermann, 2022) and being interpreted in radically different ways by a variety of actors, including social movements, civil society, and Indigenous communities (Couture and Toupin, 2019), the concept of digital sovereignty has been increasingly used by states in their policy discourses to designate a reasserting of their power in regulating relations between private corporations and individuals. This has been the case not only for authoritarian regimes such as China and Russia but also for Brazil and the EU (Pohle and Thiel, 2020), which have seen a marked rise in initiatives for state sovereignty over the digital. Of course, states – these "weary giants of power and steel" (Haggart et al., 2021) – had never fully abdicated control. And yet, in recent times of "techlash" (Weiss-Blatt, 2021), when public trust in tech corporations has significantly declined, states have become much more confident in reclaiming power over tech policy, interfering not only when it comes to data protection but also to matters of taxation, fighting hate speech, cybercrime, etcetera.

That said, the idea of state data sovereignty is by no means a simple or easy alternative to the current liberal regime of data protection. For example, the EU, one of the biggest champions of digital sovereignty, recently adopted an approach to digital sovereignty that is fully complementary with, rather than contradicting, a focus on data as a personal asset. In other contexts, the liberal respect for individual rights in data protection has been replaced, in the name of digital sovereignty, not by more collectivist socialist understandings but by more authoritarian approaches to managing populations through extensive use of big data, as seen in China, for example (Aho and Duffield, 2020). Thus, rather than proposing state digital sovereignty as a panacea and easy alternative to current inefficient regimes of data protection in surveillance capitalism, the question that emerges is rather what type of digital sovereignty we aim for. In the last section of this concluding chapter, we discuss the role of civil society actors and various visions of democratizing digital sovereignty and giving power over data "to the people."

Data Activism and Sovereignty "from Below"

Most contributions to this volume have focused on political parties, governments, and institutional state actors and their uses (and abuses) of big data. An important and overlooked dimension of artificial democracy, however, has to do with different forms of bottom-up intervention and activism, often outside the realm of institutions. Indeed, as state regulation has attempted to keep pace with developments in the digital world, civil society has also been active in this area, including media organizations, NGOs, and protest movements attempting to regain control over their data and understanding digital sovereignty not only, and not so much, as an attribute of the state but above all as social movements', Indigenous communities', and civil society's sovereignty over their data and communications (Couture and Toupin, 2019). The last decade has seen a rise in Indigenous activism, data activism, hacktivism, and a variety of citizen initiatives that aim to counter the worst excesses of datafication and commodification of citizen data by private companies, but also by governments (Data Sphere, 2022; Lehuedé, 2022; Mattoni and Odila, 2021; Milan, 2017; Milan and Treré, 2019).

Authors have distinguished between reactive and proactive data activism. Reactive data activism takes place "when citizens resist the threats to civil rights, and privacy, that derive from corporate intrusion and government surveillance. They do so primarily by means of technical fixes or by creatively subverting and hijacking the monitoring and snooping with tactics like countersurveillance and obfuscation but adopt also more traditional movement strategies such as campaigning" (Milan, 2017, 5). Proactive data activism, on the contrary, "actively takes advantage of the possibilities for advocacy and campaigning that big data offer and uses and appropriates data to foster social change. It articulates the link between the right to use data and a functioning public sphere on the grounds that access to data equals empowerment" (Milan, 2017, 6).

Another two sets of key distinctions to understand the contentious politics of data are, first, "data as stakes" (data as issues and objects of political struggle in their own right) versus "data as repertoires" (data as tools of political struggle), and second, "individual practices" versus "collective action" (Beraldo and Milan, 2019). Recent years have seen an increase in individualized forms of resisting private and governmental surveillance through the adoption of various privacy and security tools, especially in the context of activism (Horstman, 2022; Tadic et al., 2023). Such individual and largely reactive "solutions" have been combined with various forms of collective action. Social movements and NGOs such as the Electronic Frontier Foundation or the European Digital Rights Initiative, fighting for "data as stakes," have mounted numerous campaigns in defence of data privacy, combining the use of big data with more traditional campaigning strategies. Notably, hacktivists have often employed their coding skills to oppose laws attacking internet freedom (Rone, 2020) or governments perceived as authoritarian and suppressing the civil rights of their citizens (Coleman, 2012), among others. While such actions have been largely reactive, a good example of collective, proactive data activism are the citizen mobilizations to make sense of open government data and to tackle corruption, such as the Ficha Limpa (or Clean State Law) and the Ten Measures Against Corruption campaigns in Brazil (Mattoni and Odilla, 2021). Other recent examples of proactive citizen campaigns have been initiatives such as Make Amazon Pay,[1] which is fighting to "Make Amazon Pay fair wages, its taxes and for its impact on the planet," or Collectivize Facebook,[2] which has started a collective legal action to turn Facebook into collective property.

In all these cases, data activism can be interpreted as an attempt to "bring democratic agency back into the analysis of how big data affect contemporary society" (Milan, 2017, 2). Citizens, in this interpretation, are not passive victims of the “redrawing of the social contract” but are rather active participants in this process, not only defending their rights through reactive activism but also attempting to keep companies and governments in check with proactive campaigns. These types of campaigns draw attention away from the more state-centred understandings of digital sovereignty discussed above and bring attention to the collective democratic aspects of digital sovereignty.

Indeed, one way to get out of the impasse of defending state sovereignty, in the face of rising authoritarianism and the politics of emergency, is to emphasize the popular democratic foundation of digital sovereignty (Rone, 2023b), giving people more power to democratically control decision making on data collection, analytics, and use through the means of representative democracy but also through forms of counter-democracy (Rosanvallon, 2008), including protests, better journalistic oversight, and court cases (Rone, 2023a). Others have gone further, developing frameworks for community data ownership (Singh and Vipra, 2019), “data-owning democracies” (Fischli, 2022), or “platform socialism” (Muldoon, 2022). In some of these visions, data are owned collectively or publicly – often at the local or regional level but potentially also at the state level – and digital sovereignty is thus conceived at multiple levels of governance. Similar proposals have been made in terms of collective ownership, governance, and participation when it comes to decision making over campaigning data collected by parties (Rone, 2023b).

Proposals for moving beyond activism and actually institutionalizing citizen power and “giving power over data to the people” are still in their infancy, however. Ultimately, whatever form the new social contract between citizens, governments, and corporations takes (see Lyon’s contribution to this volume), when it comes to data governance, this will be the result not simply of academic reflection but also of activism, political coordination, and mobilization. The relationship between data and democracy is problematic, fluid, and highly contested and democratizing decision making over data seems to be a possible way to resolve current dilemmas and problems with data privacy breaches, manipulation, and microtargeting. The same problems, however, could also be "resolved" in radically different ways, including through tech corporations amassing even more power or states becoming even more controlling in the future, with democracy falling behind.

This volume's contributions on the ways data governance relates to policies, politics, and polity are thus crucial to opening a broader debate about the future of democracy in the age of big data, AI, and comprehensive digitization. The word *artificial* as used in the title of this book can have two main meanings: it could mean "made or produced by human beings" or "insincere, not real, fake." It is up to our collective agency to make sure that "artificial democracy" ends up meaning a democracy that we all make collectively, through deliberation and conflict alike, rather than a deep-fake, an imitation of a form of governance that we once had but is long gone and exists only as a simulation. It is up to everyone – practitioners, academics, and citizens – to take the full scope of the transformation that is taking place. And to think, speak, and debate on the future of democracy.

Notes

[1] See https://makeamazonpay.com.

[2] See https://collectivize.org.

References

Aho, B., and Duffield, R. (2020). Beyond surveillance capitalism: Privacy, regulation and big data in Europe and China. *Economy and Society*, *49*(2), 187–212.

Anstead, N. (2017). Data-driven campaigning in the 2015 United Kingdom general election. *The International Journal of Press/Politics*, *22*(3), 294–313.

Baldwin-Philippi, J. (2017). The myths of data-driven campaigning. *Political Communication*, *34*(4), 627–33.

Bauman, Z., Bigo, D., Esteves, P., Guild, E., Jabri, V., Lyon, D., and Walker, R.B. (2014). After Snowden: Rethinking the impact of surveillance. *International Political Sociology*, *8*(2), 121–44.

Beraldo, D., and Milan, S. (2019). From data politics to the contentious politics of data. *Big Data and Society*, *6*(2).

Cheeseman, M. (2022). *Web3 and communities at risk: Myths and problems with current experiments.* Minderoo Centre for Technology and Democracy, University of Cambridge.

Coleman, G. (2012). *Coding freedom: The ethics and aesthetics of hacking*. Princeton University Press.

Couture, S., and Toupin, S. (2019). What does the notion of "sovereignty" mean when referring to the digital? *New Media and Society*, *21*(10), 2305–322.

Data Sphere. (2022). *Initiatives to follow on Indigenous Data Sovereignty*. https://www.thedatasphere.org/news/initiatives-to-follow-on-indigenous-data-sovereignty/.

Dommett, K. (2020). Regulating digital campaigning: The need for precision in calls for transparency. *Policy and Internet*, *12*(4), 432–49.

Dowling, M.E. (2022). Cyber information operations: Cambridge Analytica's challenge to democratic legitimacy. *Journal of Cyber Policy*, *7*(2), 230–48.

Fischli, R. (2022). Data-owning democracy: Citizen empowerment through data ownership. *European Journal of Political Theory*, *23*(2), 204–23.

Floridi, L. (2020). The fight for digital sovereignty: What it is, and why it matters, especially for the EU. *Philosophy and Technology*, *33*, 369–78.

Haggart, B., Tusikov, N. and Aart Scholte, J. (Eds.). (2021). *Power and authority in internet governance return of the state?* Routledge.

Horstmann, N.D. (2022). The power to selectively reveal oneself: Privacy protection among hacker-activists. *Ethnos*, *87*(2), 257–74.

Kefford, G., Dommett, K., Baldwin-Philippi, J., Bannerman, S., Dobber, T., Kruschinski, S., Kruikemeier, S., and Rzepecki, E. (2022). Data-driven campaigning and democratic disruption: Evidence from six advanced democracies. *Party Politics*, *29*(3), 448–62.

Lambach, D., and Oppermann, K. (2022). Narratives of digital sovereignty in German political discourse. *Governance: An International Journal of Policy, Administration and Institutions*, *36*(3).

Lehuedé, S. (2022). Territories of data: Ontological divergences in the growth of data infrastructure. *Tapuya: Latin American Science, Technology and Society*, *5*(1), 2035936.

Manokha, I. (2018). Surveillance: The DNA of platform capital – the case of Cambridge Analytica put into perspective. *Theory and Event*, *21*(4), 891–913.

Mattoni, A., and Odilla, F. (2021). Digital media, activism, and social movements' outcomes in the policy arena: The case of two anti-corruption mobilizations in Brazil. *Partecipazione e conflitto*, *14*(3), 1127–150.

Milan, S. (2017). Data activism as the new frontier of media activism. In V. Pickard and G. Yang (Eds.), *Media activism in the digital age* (151–63). Routledge.

Milan, S., and Treré, E. (2019). Big data from the south(s): Beyond data universalism. *Television and New Media*, *20*(4), 319–35.

Muldoon, J. (2022). *Platform socialism: How to reclaim our digital future from big tech*. Pluto Press.

Pohle J., and Thiel, T. (2020). Digital sovereignty. *Internet Policy Review*, *9*(4), 1–19.

Rone, J. (2020). *Contesting austerity and free trade in the EU: Protest diffusion in complex media and political arenas*. Routledge.

Rone, J. (2023a). The shape of the cloud: Contesting date centre construction in North Holland. *New Media and Society*, online first. https://doi.org/10.1177/14614448221145928.

Rone, J. (2023b). Beyond Brexit? Public participation in decision-making on campaign data during and after referendum campaigns. *Media and Communication*, *11*(1), 69–80.

Rosanvallon, P. (2008). *Counter-democracy: Politics in an age of distrust*. Cambridge University Press.

Simon, F.M. (2019). "We power democracy": Exploring the promises of the political data analytics industry. *The Information Society*, *35*(3), 158–69.

Singh, P.J., and Vipra, J. (2019). Economic rights over data: A framework for community data ownership. *Development*, *62*(1–4), 53–57.

Solove, D.J. (2013). Introduction: Privacy self-management and the consent dilemma. *Harvard Law Review*, *126*(7), 1880–903.

Stromer-Galley, J. (Ed.). (2019). Introduction: The paradox of digital campaigning in a democracy. In *Presidential campaigning in the internet age* (14–20). Oxford University Press.

Tadic, B., Rohde, M., Randall, D., and Wulf, V. (2023). Design evolution of a tool for privacy and security protection for activists online: Cyberactivist. *International Journal of Human–Computer Interaction*, *39*(1), 249–71.

Tufekci, Z. (2014). Engineering the public: Big data, surveillance and computational politics. *First Monday*, *19*(7).

Weiss-Blat, N. (2021). *The techlash and the tech crisis of communication*. Emerald Publishing.

Zuboff, S. (2019). *The age of surveillance capitalism.* Profile Books.

Zuiderveen Borgesius, F., Möller, J., Kruikemeier, S., Ó Fathaigh, R., Irion, K., Dobber, T., Bodo, B., and de Vreese, C.H. (2018). Online political microtargeting: Promises and threats for democracy. *Utrecht Law Review*, *14*(1), 82–96.

Contributors

Colin J. Bennett is professor emeritus of political science and Fellow at the Centre for Global Studies at the University of Victoria, British Columbia. For over thirty years, his research has focused on the comparative analysis of privacy policy at domestic and international levels. In addition to numerous scholarly and newspaper articles, he has published seven books on these subjects, including *The Governance of Privacy* (MIT Press, 2006), as well several policy reports for national and international agencies. His current work focuses on the importance of privacy for democratic rights and on the capture and use of voters' personal data by political parties in Western democracies.

Cecilia Biancalana is an assistant professor in the Department of Culture, Politics and Society at the University of Turin. Her research focuses on political ecology, party change, populism, and the relationship between the internet and politics.

François Blais has been a professor of political theory at Laval University since 1992. He was also a minister in the Government of Quebec from 2014 to 2018. His main areas of interest are distributive theories of social justice and theories of democracy.

Pierre-Luc Déziel is a full professor of privacy law at the Faculty of Law, Université Laval, and holder of the Canada Research Chair in Health Data Protection and Valorisation. His main research interest is the impact of new technology of privacy rights in the health context.

Yannick Dufresne is an associate professor in the Department of Political Science at Université Laval, holder of the Chair in Leadership in Teaching Digital Social Sciences (CLESSN), and director of the Centre for Public Policy Analysis. He specializes in public opinion and in measuring attitudes on complex political issues. He has also contributed to the design and development of several large-scale digital data collection applications such as the Vote Compass, Project Quorum, Datagotchi, and the Polimeter.

David Lyon is former director of the Surveillance Studies Centre at Queen's University, and professor emeritus of sociology and law. He has authored, co-authored, or edited thirty-three books. His new book is *Surveillance: A Very Short Introduction* (Oxford, 2024) and his current research project is a book for Polity Press, *Surveillance and the Eye of God: Medieval Roots and Modern Shoots.*

Eric Montigny is professor of political science at Université Laval. He is codirector of the associated international laboratory on political parties, representation, and sustainable development. His research focuses on political parties, electioneering, political cleavages, democracy, and parliamentarism.

Catherine Ouellet is an assistant professor in the Department of Political Science at the Université de Montréal. She specializes in Quebec and Canadian politics, focusing her research primarily on lifestyle, public opinion, and political behaviours in Quebec and Canada. In recent years, she has gained rigorous training in digital social sciences methodologies and approaches. Notably, she is the cocreator of Datagotchi, a playful data collection tool. Her research has been published in various journals, including *Nations and Nationalism*, *International Journal of Public Opinion Research*, *Parliamentary Affairs*, *French Politics*, and the *Canadian Journal of Political Science*.

François Pellegrini is professor of informatics at Université de Bordeaux, where he also teaches digital law. He is a researcher at LaBRI and Inria, specializing in high-performance parallel computing, graph algorithms, software law, and personal data law. He is a former vice-president of the French Data Protection Authority (CNIL).

Julia Rone is a postdoctoral researcher at the Minderoo Centre for Technology and Democracy at CRASSH at the University of Cambridge. She was previously a Wiener-Anspach postdoc at the Université libre de Bruxelles and the Department of Politics and International Studies, Cambridge. Her current research focuses on the democratization of internet policy regulation. She has written on hacktivism, digital disobedience, and more recently, the rise of far-right media in Europe.

Pierre Trudel is an emeritus law professor at Public Law Research Centre, University of Montreal. From 2018 to 2020, he was a member of the Expert Panel for the Revision of Canadian Broadcasting and Telecommunications Laws. He has published numerous books on information technology law. He is a regular columnist for *Le Devoir*. https://pierretrudel.openum.ca/.

Index

Note: “(f)” after a page number indicates a figure; “(t)” after a page number indicates a table.

This page intentionally left blank

Printed and bound in Canada

Set in Myriad, Times New Roman, and
Sabon by Artegraphica Design Co.

Copy editor: Candida Hadley

Proofreader: Carmen Tiampo

Indexer: Stephen Ullstrom

Cover designer: Michel Vrana

Authorized Representative:
Easy Access System Europe
Mustamäe tee 50, 10621 Tallinn, Estonia,
gpsr.requests@easproject.com

This page intentionally left blank