Centre for International
Governance Innovation

CIGI Papers No. 160 — February 2018

A National Data Strategy for Canada

Key Elements and Policy Considerations

Centre for International
Governance Innovation

CIGI Papers No. 160 — February 2018

A National Data Strategy for Canada

Key Elements and Policy Considerations

CIGI Masthead

Printed in Canada on paper containing 10% post-consumer fibre and certified by the Forest Stewardship Council® and the Sustainable Forestry Initiative.

Centre for International Governance Innovation

67 Erb Street West
Waterloo, ON, Canada N2L 6C2
www.cigionline.org

Table of Contents

Executive Summary

This paper outlines the key elements of a data strategy for Canada. It has been prepared under the overall coordination of CIGI President Rohinton P. Medhora, with the input of several leading scholars and practitioners.[1] The paper starts by making the case for a data strategy and describing its possible ambit. It considers the sectors in Canada where a national data strategy is most needed and examines the domestic and international policy considerations. While this paper cannot flesh out such a strategy in detail, it is hoped that it will start a much-needed conversation among political leaders, policy developers, business people and civil society about how Canadians can prosper from the data revolution in a way that respects our fundamental values.

The Importance of Data and the Need for a National Data Strategy

Canadians are coming to the realization that the data revolution has great economic potential. Indeed, some have already hailed data "the new oil."[2] This may be an imperfect analogy, but it does capture the excitement and high expectations surrounding the data-driven economy. The prospect of extracting lucrative insights from rapidly growing pools of data is galvanizing entrepreneurs and investors in all sections of industry.

There is no doubt that ownership of data and associated analytical algorithms has taken on great importance for the future of many, if not all, commercial enterprises. The success of the most valuable companies in the world (Apple, Google, Facebook and Microsoft) is now underpinned by, above all else, a sophisticated capacity to collect, organize, control and commercialize stores of data and intellectual property (IP). Big data and artificial intelligence (AI) are fast becoming the leading drivers of wealth creation, and are increasing productivity, accelerating innovation and disrupting existing business models. All companies will soon become data and IP companies. John Deere, for example, no longer simply manufactures tractors — it now also collects data on the farms where those tractors are used. The company plans to leverage this data in the coming years to shift the control and profit structure of farming, similar to how Uber upended the taxi industry.

But there is an equally if not more important non-economic dimension to the data revolution. In our rush to profit from data, we must be sensitive to the fact that it is not a commodity like grain or timber. Once created, data — and especially personally identifiable information — exercises an enduring and uniquely potent influence on individual lives, social relationships and autonomy. While there is still debate about whether individuals "own" the data that relates to them, it is undeniable that they retain a stake in that data — who sees it and how it is used. Finding ways to respect this interest while commercializing the data will be a central mandate of any data strategy.

More broadly, we have seen how a greater capacity to access and manipulate data can alter our political landscape. *The Guardian* has exposed the vulnerability of democracies by showing how Mercer and Cambridge Analytica deployed a data strategy on platforms such as Facebook to influence the outcomes of both the Brexit referendum and the 2016 US presidential race (Cadwalladr 2017). More recently, *The Washington Post* has detailed Russian use of data-driven Facebook messaging campaigns to affect the outcome of US elections (Dwoskin, Timberg and Entous 2017). In short, the data revolution not only has huge implications for commerce, but for the very operation of liberal democracy itself (Owen 2017).

Any data strategy will have to address both the economic and non-economic dimensions of harnessing big data. Balances will have to be struck between numerous goals:

- reaping the gains from the economic potential of data;

1 They are: Sachin Aggarwal (Think Research); Anil Arora (Statistics Canada); Lisa Austin (University of Toronto); Dan Breznitz (University of Toronto); Dan Ciuriak (Ciuriak Consulting); Andrew Clement (University of Toronto); Ian MacGregor (NW Refining); Kurtis McBride (Miovision); Jonathan Obar (York University); Teresa Scassa (University of Ottawa); and Paul Vallee (Pythian).

2 See www.quora.com/Who-should-get-credit-for-the-quote-data-is-the-new-oil.

- respecting, or even enhancing, its fundamental privacy elements;
- preserving an open society and democracy;
- maintaining public security; and
- building institutions (such as information networks and governance processes) that maintain or enhance Canada's national identity.

The delicate interplay between these goals means that they should be addressed together, within a single strategic framework. Given the speed at which the data revolution is moving, there is an urgent need for such a framework.

Data Economics: Valuing Canada's Existing and Future Stocks and Flows of Data

Data is the essential capital stock of the data-driven economy, just as IP is the essential capital stock of the knowledge-based economy and physical capital is the essential capital stock of the industrial economy. Data is not acquired through transactions that leave a paper trail of payments and receipts — hence, data is nowhere to be seen in traditional national economic accounts, international trade statistics or the quantitative cost-benefit analysis of government policies. The value of data is, however, indirectly perceived in the valuation markets place on the intangible assets of data-driven firms such as Google (current market cap of approximately US$14 billion) Facebook (approximately US$517 billion) and Uber (US$50 billion or more). This puts the likely market value of data in the trillions of dollars at the dawn of the data-driven-economy era, with potential for even greater expansion as the digital transformation races forward.

Of course, data first needs to be captured, and capture is privileged to those controlling the infrastructure or the apps. This results in powerful network externalities that drive "winner-take-most" market structures (Google market cap: US$580 billion; Yahoo: US$4.5 billion). Growing Canada's share of global data capital — and therefore broadening its stake in the global data-driven economy — will depend heavily on the extent to which Canadian firms can capture data in areas where first movers have not already established insuperable dominance. Industrial and innovation policy for the data-driven economy will thus hinge on how to best capitalize on Canada's proprietary data based on privileged access: in some instances, private sector-led commercialization may be optimal; in others, it might be better to make data available to the public at sub-market prices.

Above all, decision makers need to recognize and account for the potential use and exchange value of this data when developing and implementing policy. As it stands, Canada is negotiating international treaties in which it effectively gives away its data — the key capital asset of the digital age. Governments are making Canadian data freely available internationally (for example, by committing to the free flow of data across borders, and by foregoing data localization requirements), and are seemingly oblivious to the unfavourable value proposition this represents. It is, of course, imperative to many businesses that certain data sets be shareable across borders, but the current balance needs to be recalibrated to recognize that Canadian data has value. A data strategy must facilitate free flows of information to be sure, but not "give away the store" to international competitors.

Technical Infrastructure for a Data-enabled Society

Canada's national data strategy should embrace the benefits of distributed computing. Too many data strategies are still premised on a mainframe-based architecture, where all decisions are centralized, creating a massive load on existing IT systems. According to Anthony Townsend (2014), distributed computing will come to define the management of urban spaces: "Smart cities are going to look much more like the web, where there's going to be a lot of things deployed by

individual decision, talking to each other through open standards in very ad hoc, loosely knit ways." Townsend's Palo Alto-based Institute for the Future espouses this smart city model, aimed at developing a "complex and vibrant system instead of a controlled, hierarchical one" (McBride 2015).

Canada needs to ensure that this sort of distributed-system architecture serves as the model for future data collection infrastructure. A distributed network of connected sensors will enable the kind of data harvesting needed to power broader sets of transformative applications. This decentralized, easy-to-access mesh of sensors will be everywhere: from farm fields to factories, city intersections to electrical grids, and health monitors to smartphones. Canada's timely drive to bridge the remaining digital divides means that the entire country — from rural and remote communities to large cities — will soon be well-positioned to make this happen.

An indispensable attribute for such a system will be interconnectedness. Here, our existing telecommunications infrastructure offers opportunities, but also constraints. The mesh will not only leverage key wireless standards for the cellular, public and industry-regulated spectrum, but will incorporate open standards for device and data access. At the same time, it will have to work around legacy designs. All data exchange will nonetheless require authenticated access, ensuring all data is secure and adheres to privacy legislation.

A no-less-important attribute of the mesh will be communal access and opportunities for collaboration among Canada's innovation ecosystems. The mesh should be available on a reasonable basis to any developer who wants to build applications or gain access to data. Integration with the network being built out for the Internet of Things (IoT) will mean that hundreds of other parties will be able to create new processes and intelligence via data application programming interfaces (APIs).

Key Sectors Where Canada Needs a National Data Strategy

Energy, Mining and Agriculture

The realm of consumer data is already dominated by Google, Amazon and Facebook, which have been collecting and cataloging data for a decade. Surprisingly, however, no analogue exists in the resource and agricultural sectors. On the rare occasions that useful machine-learning data is collected in these sectors, it is fragmented, siloed and proprietary. A few actors collect the type and amount of data that is needed for machine-learning applications, but only for their own narrow, in-house purposes.

Canada is uniquely positioned to become a world leader in the application of big data and machine learning to primary industries. Its advantage lies in the relative scale and modernity of the resource harvesting and processing operations. Canadian agriculture, energy production, mining and forestry are generally equipped with the modern, well-instrumented infrastructure necessary for collecting the types and quantities of data that are required for the application of machine-learning methods. The instruments exist, but their potential needs to be fully leveraged. Just as Uber and Airbnb identified and capitalized on unutilized capacity in vehicle and housing stocks, Canadian companies should seek ways to generate data — and valuable knowledge — from the trucks, drills, conveyors and processors that are already out in the field. Outfitting these implements with the necessary sensors, or connecting to the ones they already have, will enable the data harvesting necessary to make Canada the leader in resource industry and agriculture machine learning. We do not have the market to be a global leader in self-driving car data, but we have both the market and the expertise to lead in machine learning for resource and agriculture industries.

This could not come at a better time for companies in Canada's primary industries, which are struggling to manage costs in a highly competitive environment. The significant insights and efficiencies that come from the application of machine learning will help make transformative improvements in the financial and environmental performance of these enterprises.

While the benefits of machine learning are often obvious to corporate management, the myopic views of shareholders seldom allow for the long-term investments that are needed to implement it. Google is now a dominant player in the self-driving car industry, but it took a decade to collect the data to make that happen. It will take similar resolve and a long-term outlook from Canadian investors to collect the data upon which machine learning works. A further impediment is that these companies typically do not have the academic and scientific expertise to identify the best opportunities for data collection and machine-learning applications. The players in these industries are waiting for leadership on developing the data infrastructure, together with the academic and commercial talent to utilize it.

A deep, broad, open-source data library of all types of resource and agriculture industry data is an essential requirement for Canada's globally competitive industries. By building this basic data infrastructure and making it a common resource for all, the competitiveness of these industries can be enhanced, transformational improvements in their environmental performance can be made and lasting opportunities can be provided for young Canadians. No single company or industry can or should have a monopoly on the collection and control of this information — its value can only be maximized for Canada through public ownership. Data is the twenty-first-century equivalent of the railway, and we do not want to see it in private hands this time.

Health Care

The volume of digital health information is growing exponentially (48 percent per year) and is estimated to reach over 2,000 exabytes by 2020. This information is being generated by a growing number of sources, from physicians manually logging patient data into health databases to automated sensors collecting biotelemetry. This wellspring of information is transforming how medical knowledge is created and, in turn, how health care is provided. Conventional academic approaches to research involve slow and constrained data capture and analysis. These methods do not offer the volume, depth and breadth of information required to keep pace with evolving health-care needs. Discoveries gleaned from the analysis of big data sources will eventually outpace those achieved through conventional research by 1000:1. Unlocking the potential value of health-care data will depend on the implementation of new policies, standards and technologies to facilitate open, structured and secure data sharing.

Today, vast amounts of health data remain isolated and underutilized. Health information systems in Canada — from electronic medical records to provincial data repositories to personal health platforms — lack the common standards that allow for even basic interoperability. The inability to fully share, access and understand health information across these siloed repositories prevents the development of truly comprehensive analysis of patient medical histories and the operation of the health-care system as a whole. The benefits of more data interchange include:

- increasing the operational efficiency of care delivery;
- better monitoring of emerging epidemiological trends;
- improved clinical decision making and risk management;
- delivery of more effective personalized medicine;
- enabling the application of AI and machine learning for continuous improvement of care algorithms and standards; and
- accelerating medical research.

Investments that facilitate access, manipulation and analysis of health-care data assets will also generate large amounts of commercial IP. Ultimately, those that own large parts of the medical information life cycle, from data capture to analysis, will be the economic winners.

Ubiquitous access to health information will have a transformational impact on patients, clinicians and the health -care system at all levels. Health care today is organized as a hierarchy of different service providers, from specialists to personal support workers, all practising in diverse care settings. Patient care is most expensive when delivered from the top of the pyramid by hospital specialists, yet this is how most care is currently delivered. The hierarchy is largely the result of asymmetrical access to health-care knowledge for those at the top. The democratization of knowledge through open data will flatten the hierarchy by disseminating information more evenly, allowing those at the bottom to take on more work of the higher-tier specialists. This will gradually move more care into the local community setting and decrease overall costs.

Increased access to health data will also empower patients to play a greater role in their own care, allow policy makers to make better decisions and help pharmaceutical and biotech companies develop better products more efficiently. Canada requires a national strategy to unlock the latent value in health-care data. Doing so will reinforce the country's role as a global leader in health care innovation.

Cities

Cities need to think beyond open data initiatives when planning their technology strategy. Open data represents the last mile, which allows cities to share a fraction of their data via a portal with innovators, citizens or researchers. Cities should focus first on adopting truly open technology. If an important part of the technology stack is not open, standards-based and API-driven, then cities will struggle to reap the full benefits of an open ecosystem, no matter how many open data initiatives are undertaken.

Unless they commit to building their infrastructure with open technology, cities will struggle to fulfill their vision for a smart urban environment. Open technology will empower cities to plan, grow and innovate on their own terms, and without being beholden to a small number of vendors. This represents a departure from the closed technology investments that have shaped city infrastructure and planning for decades Embracing open technology can lead to:

- more sustainable and stable infrastructure investments;
- more efficient procurement;
- more options for development by avoiding "vendor lock-in";
- greater cross-departmental collaboration;
- greater civic transparency; and
- conducive conditions for a vibrant innovation ecosystem.

The elements of open technology that cities should seek in every infrastructure investment are listed below. They should be part of a city's open technology mandate:

- standards-based design (no use of proprietary protocols);
- open APIs to provide access to all raw and processed data (programmatic access to all data);
- technical disclosure of all systems (i.e., no "black box mystery technology");
- use of third-party certification for compliance with privacy and security best practices; and
- data portability to allow export/import to other systems.

The foregoing principles will ensure that cities have control of their own technology stack, the ability to integrate it into all other systems and the ability to innovate on top of their infrastructure (themselves or via other partners). It also ensures they can share their data with citizens, researchers and other stakeholders.

Modernizing, Streamlining and Opening Government Data

Fuelled by an explosion in data generation (for example, sensors, personal devices, and administrative and regulatory data), big data and data providers, governments and citizens alike are seeing rapid advancements in enabling technologies and analytical tools, as well as the emergence of new players and innovative partnerships and business models.

Despite significant new investments and historic approaches to innovation, the Government of Canada's approach to the data revolution is largely piecemeal and reactive. All federal departments and agencies collect, generate and use data, whether it is in the form of tax returns, social programs, research funding, fuel consumption statistics, health data, immigration flows, weather information, geospatial maps of buildings or land, crop inventories or business program activities. Across the federal family, data is not consistently managed as a strategic asset, programs are not targeted and IP regimes are not sufficiently mature to drive data as a strategic or competitive advantage. Departments are procuring and holding the same data sets, meaning

the government may be paying for the same asset multiple times, and efforts to make data truly open and accessible are not coordinated. Departments and agencies are, therefore, unable to maximize their ability to achieve outcomes, seek efficiencies and ensure that we are not simply getting data rich and, at the same time, knowledge poor.

Complementary and coordinated initiatives led by departments and agencies (such as Statistics Canada, the Competition Bureau, the Treasury Board Secretariat, Shared Services Canada, the Canadian Standards Council and the National Research Council) and woven together will provide the necessary framework to propel the vision and strategy. Statistics Canada is a key, but underacknowledged, player in the data field, and is well positioned to enable and support a government-wide move to a data-driven future. Data science expertise, analytical infrastructure and strong privacy controls are key aspects of the role Statistics Canada already plays. Building on its reputation as a world-leading independent agency, with nearly 100 years of data management experience, Statistics Canada could play an important role.

A comprehensive strategy for a whole-of-government approach to data could be comprised of three elements:

- Increasing access to data to drive innovation and inclusion: Investments are needed to harvest real-time data through the myriad of sources and channels available today. Similarly, investments in disseminating online anonymized micro-data and provisioning big data with the appropriate analytical and visualization tools are essential. To counter the plethora of suspect data and poor-quality data on the internet, the Government of Canada needs to play a far more prominent role in providing unbiased and well-documented data and information through multiple channels. This could be achieved via significant improvements to the Statistics Canada websites, enhanced government-wide web portals and the development of leading-edge data visualization to promote uptake across the spectrum of data expertise.
- Mobilizing data: Investments in robust, scalable and modern technical, statistical and legal infrastructure are essential for the Government of Canada to deliver on its role as an effective data steward. There is a need to break down the policy, legislative, cultural and technological barriers built by outdated laws and regulations between departments and other levels of government to open up federal data holdings. These include the Personal Information Protection and Electronic Documents Act, the Statistics Act and the Privacy Act, all of which are currently under review or being considered for modernization. Key investments would enable multiplying the value of data holdings, while safeguarding privacy and confidentiality.
- Increasing data literacy and statistical capacity building: Initiatives are required to coordinate and build data literacy and numeracy across the school and university systems, as well as in governments and businesses. These could seek to increase the data usability, relevance, knowledge and products for all Canadians, from kindergarten children to post-doctoral candidates, as well as from public servant generalists to expert data-leveraging entrepreneurs. Government action here could enable new opportunities for the next generation of Canadians and public servants to develop and create value from data, by ensuring the right skill sets and capacities exist and are nurtured, and that they increase innovative uses of data and information.

Open Data

The federal government has been quite active around open data — it has done work on standards and has tried to encourage uptake of a more or less standard licence across the country to facilitate interoperability. It maintains an open data portal and has projects in place to expand the amount of available open data.

Open data in municipal and state governments is becoming more common.[3] The definition of "open," however, is not 100 percent consistent. The most widely adopted definition is perhaps the International Open Data Charter, which gives the six principles of open data as:[4]

- open by default;
- timely and comprehensive access to data;

3 See index of open data portals worldwide: http://dataportals.org.

4 See https://opendatacharter.net/.

- accessible and usable data;
- comparable and interoperable;
- data for improved governance and citizen engagement; and
- data for inclusive development and innovation.

The group provides a charter for adoption by cities/states.[5] In Canada, it has been adopted by Edmonton and the Province of Ontario.

Specific open data standards have emerged to help cities make certain types of public data open and interoperable. Examples include:

- The Toronto-based World Council on City Data (WCCD), which hosts a network of innovative cities committed to improving services and quality of life with open city data, and provides a consistent and comprehensive platform for standardized urban metrics. The WCCD is implementing ISO 37120 Sustainable Development of Communities — Indicators for City Services and Quality of Life,[6] the new international standard created by cities, for cities. The WCCD has developed the first ISO 37120 certification system and the Global Cities Registry™.[7]
- The General Transit Feed Specification, an open standard for transit data, developed as a side project within Google.[8]

These standards are making it easier to transition specific parts of government to open data when there is cooperation from industry.

There are particular challenges to developing open data in a federal state. Developing standards and licensing for adoption by all levels of government in order to facilitate interoperability is an imperative. Also, it might be good to develop strategies to encourage government workers to use open data not just within their own level of government (i.e., combining data sets from different departments at the federal level to obtain new insights), but to also encourage the use of data from all levels of government where relevant.

"Real-time" data presents new and different challenges. In a national strategy, some consideration must be given to new forms of open data such as real-time data, which will likely need its own standards, policies and licences.

Creating Successful Innovation Outcomes for the Canadian Economy

Canada's innovation ecosystem must be underpinned by a carefully designed open architecture technology strategy that includes, but goes well beyond, open data. Applications using open data are growing, but they remain largely niche and one-way (i.e., taking open data and building some visualization, reporting or analysis on top). As with cities, if the rest of the technology stack is not open, standards-based and API-driven, then any strategy will struggle to access the true potential of open data.

Open technology is focused on empowering participants to plan, grow and innovate on their own terms, and not be limited or trapped by a vendor. There is a critical ongoing role for the Standards Council of Canada in the implementation of all of these strategies.

It will also require a strategy for smart sensors and interfaces (the IoT). Every process, touchpoint and piece of infrastructure will undergo a renewal in the decade ahead. It is a major disruption of the old approaches, and Canada can drive its economic productivity and prosperity if an open architecture technology strategy is used to successfully integrate real-time workflows.

But it cannot wait. The very large base of devices that are installed today should be used to immediately start building the "self-driving" data that will lead to machine-learning opportunities. The existing base of sensors and instruments in all of Canada's key sectors can provide the bridge to getting the needed data while the standards for new devices and methods are being figured out. There are

5 See http://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter-charter_F.pdf.

6 See https://www.iso.org/obp/ui/#iso:std:iso:37120:ed-1:v1:en.

7 The Global Cities Registry is an internationally recognized list of cities that meet state of the art requirements on the production, management and use of city data. See www.dataforcities.org/global-cities-registry.

8 See https://developers.google.com/transit/gtfs/.

hundreds of thousands of places where there is an existing data source, and the most promising must be identified and collecting and cataloguing must get under way. This immediate activity can inform the developing standards part of the process that will lead to true Canadian leadership in this area.

One of the best examples of a successful open-architecture strategy is the below-ground and production data that is a regulatory requirement for the oil and gas industry in Alberta. All drillers in the province are required to catalogue their wells in a standardized way and collect and submit geological and production data gathered during the lifetime of the well. This data is accessible to all and has been a basis for generations of energy innovation, technology development and entrepreneurial activity in Alberta. If the infrastructure and regulatory environment enables this open-architecture approach, the benefits of machine learning and the opportunities will be as plentiful for other Canadians as they were for Albertans. Capturing the opportunity and doing it before someone else figures it out is critical. Canada is dominant in a number of basic industries, it must use that position and expertise to empower the next generation. All they need is the data. Building machine-learning data collection and reporting requirements into regulatory licensing schemes would be a good start.

Care will have to be taken, however, to ensure that the economic outcomes of the data-driven economy lead to positive social outcomes. Evidence is slowly emerging to suggest that the data-driven economy is not always a driver of inclusive growth and can, in fact, exacerbate existing inequalities. Strategies to grow and promote the data-driven economy will have to include mechanisms to ensure that the new economic opportunities are available for all.

Data and Privacy Regulations

A national data strategy must not only look at ways to modernize Canada's technical data standards and infrastructure, but also at the legal frameworks governing it. The strategy should have as its core reference points the liberal democratic values that define our society. The commercial dividends of big data and AI must be realized in a way that respects and reinforces our deeper commitments to democracy, civil liberties and equality.

This does not mean that robust privacy protections need come at the expense of data-driven prosperity. Rather, the two should be seen and pursued as complementary goals. After all, a steady flow of data — the lifeblood of the data-driven economy — depends on citizens trusting that third parties will use their information responsibly. Assuring individuals of the security of their personal information and their continued autonomy in a hyper-monitored world will be key to maximizing the commercial benefits of big data. Canada's national data strategy must begin drafting the terms of a new social contract between the citizen providers of data and those in industry and government who want to collect and use it.

Canadians will not only need to be assured that private actors are handling their data appropriately, but that public sector actors — namely security and law enforcement agencies — are likewise bound by new and enhanced standards. Unbridled government access to personal information collected in a commercial context will also compromise the willingness of individuals to participate in the big data economy. As security agencies gain greater capacity to glean revealing information from the consolidation and association of disparate bits of metadata, Canadians will have to re-examine increasingly dated information-sharing practices and protocols within government bureaucracies. Statutory frameworks originally designed to regulate and protect pension and tax data are proving inadequate when it comes to regulating activities such as the surreptitious mass collection of communications metadata by security and intelligence agencies.

As data-driven decision making becomes more consequential for individuals, the remedies available for misuse of personal information must become more accessible and meaningful. New protocols and technical applications will be needed to make commitments to transparency and consent commensurate with the enormous power that data will have over peoples' lives. The old "informed consent" model, exemplified by the perfunctory pop-up, will no longer be enough. Canadians will not only need to be better informed about how their information is collected and used, but also given the tools to control the degree of their engagement with the data-driven economy.

Canadians will have to turn their attention to the international plane as well. A national data strategy will need to grapple with the fact that Canada's largest trading partner maintains markedly low privacy standards for the personal information of foreign citizens. Heidi Bohaker et al. (2015, 24) concluded the following:

- "US authorities can access Canadian persons' communications data within US jurisdiction on statutory standards that are lower than those that apply within Canada and would be unconstitutional if applied within Canada."
- "US constitutional law does not apply when US authorities access Canadian persons' communications data within US jurisdiction as long as the Canadian person remains outside of the US."
- "Even if US constitutional law did apply, Canadian constitutional law offers more privacy protection to communications data."

This unsettling approach was most recently exemplified by US Executive Order 13,768, "Enhancing Public Safety in the Interior of the United States," which states that "Agencies shall, to the extent consistent with applicable law, ensure that their privacy policies exclude persons who are not US citizens or lawful permanent residents from the protections of the Privacy Act regarding personally identifiable information" (The White House 2017). In March 2017, the United States passed a law allowing telecommunications companies to sell the browsing history of individuals and households, information that had previously been considered private citizen data (Farand 2017).

In light of these developments, Canada will need to decide how it intends to treat browsing data and whether it will ever be allowed to be transferred to foreign jurisdictions. Allowing this data to be sold overseas could mean that Canadians' browsing history might ultimately be a point of conversation at a border crossing. Similar assessments will have to be made with respect to health-care data, which will be of increasing interest to immigration officials who may wish to use it to inform border entry and immigration decisions.

Although reformed regulatory frameworks will play an important role, they will inevitably be one step behind the rapid technological changes taking place in the digital realm. Legislative action must be complemented by a broader culture shift, one that perhaps draws on the "privacy by design" mantra articulated by former Ontario Privacy Commissioner Ann Cavoukian. The success of a new social contract governing the exchange of personal information will depend on private sector and government actors incorporating privacy protections into their systems as a matter of course, rather than as a matter of compliance. Such a social contract would, in turn, lead to a new legal framework, as the current one is ill-suited to a full treatment of privacy in the age of big data.

Cyber Security, Public Safety and Civil Liberties

Canada's approach to cyber security needs to move beyond old concepts based on prevention and protecting the perimeter. While both should be secure, the strategic focus needs to shift from the network being the unit of protection to data being the unit of protection.

Edward Snowden and others have revealed the capacity for various national security agencies to infiltrate our information infrastructures. Other disclosures have shown how they hoard software vulnerabilities and deliberately weaken encryption standards for their own use rather than trying to fix them. These activities have the potential to irreparably undermine the confidence that both individuals and businesses need to have in the internet to take full advantage of its social, economic, political and cultural promise.

An urgent priority will be building a robust technical and governance regime to secure confidence in Canada's vital information infrastructures. Promoting national data/network sovereignty within democratic norms is currently vital to achieving such a regime. Carefully structured data localization, for both storage and routing, is a key element. Data localization aligns not only with individuals' interests in cyber security, but with the promotion of domestic data-driven industries. Privacy advocates and domestic innovators have an opportunity to join forces and strengthen the case for policies that advance data/network sovereignty.

The main argument against data localization is that it flies in the vision for an open, borderless, global internet, and will lead to its fragmentation — to balkanization or a "splinternet." Localization has come to be associated with repressive regimes, such as Russia and China, which favour localization to facilitate censorship and internal political control. Nationalistic anti-globalization movements in Western countries might also come to adopt data localization as part of their policy agendas. The prospect of unsavoury bedfellows should not lead us to dismiss localization entirely. Data localization is not itself the cause of political repression, nor does it necessarily represent the endorsement of an isolationist or xenophobic outlook. Maintaining an open and seamless internet is of course an important goal, but it should not be seen as threatened by an appropriately balanced and nuanced approach to localization of sensitive data. Without a degree of localization, many small countries might find themselves totally deprived of the commercial benefits of their data.

Data localization is not a panacea. The security benefits of localization efforts can be rendered less effective by offshoring, as a recent report on a Canadian bank suggested (Johnson 2017). Enhanced transparency about companies' international business relationships and offshoring policies will be required to give consumers clear choices about whom they can trust with their data.

The European Union will implement the General Data Protection Regulations (GDPR) starting in May 2018. Supporting the GDPR regulations as a minimum in Canada will make Canada a safe haven for European data, keep Canadian data sovereign from absorption into the United States and help Canadian innovators competitively. The GDPR enables separate and distinct emphasis on personally identifiable information, such as metadata, allows de-identification and delineates no-go zones.

Competition Policy and Big Data

Left unchecked, big data can and will undermine competition and harm consumers. Updated regulatory approaches are therefore critical (*The Economist* 2017). Multi-sided business models and platforms create unique challenges in assessing anti-competitive practices. Canada's Competition Bureau has an important role to play here and must be equipped with the right expertise and authority to operate effectively in the data-driven economy.

First and foremost, regulators will have to adopt a new understanding about what attributes make an enterprise a potential threat to competition. The size of a company is no longer the only, nor necessarily the most appropriate, factor to consider. Data holdings are now a key element of a business' ability to dominate a market and stifle competition. This reality must increasingly inform decision making in the regulation of competitive markets.

Another avenue to explore will be loosening the control companies currently have over data. Greater transparency from businesses about what data they collect and how they profit from it would enable policy makers and the public to better understand and respond to the new competition dynamics of the data-driven economy. Empowering individuals to retain some control over the data they generate, perhaps by allowing them to reclaim and shift their personal information among different data users, could also stimulate healthy competition. Authorities might go so far as to mandate the inter-firm sharing of certain information. Finally, governments should consider collecting and providing certain data sets as a public good, ensuring that new data-driven start-ups still have a chance to emerge.

Intellectual Property and Data

Canada's national data strategy must include an IP strategy for data. The current IP regime does not provide for ownership rights in data, although compilations of data can be protected in the North American Free Trade Agreement (NAFTA) under copyright; data can also be protected as confidential commercial information. Most often, algorithms and data are protected as trade secrets or confidential information; algorithms can also be protected under patent law. Contract law is also commonly used to govern relationships around data and algorithms, including limits on access and use.

Companies operating in the big data environment rely on a combination of different types of protection for their data.

The European Union has floated the idea of creating a *sui generis* data ownership right, and some research is being carried out in Europe to explore this concept. While a data ownership right may not be at all in Canada's best interests, it should be carrying out research on this issue in order to prepare itself to take a position on it.

While it is important to ensure some protection for valuable data assets, at the same time, it is important not to overprotect data. Copyright law has always kept "facts" in the public domain because facts are the building blocks of creativity and innovation. Excessive property rights will stifle innovation.

In addition, given the importance of data to decision making, democratic participation, citizen engagement and so on, any protection of data must come with balanced rights to access, and use data in order to protect the public interest.

Where IP issues have been on the table in international trade negotiations, this has reliably meant one thing: pressure to increase levels of protection (maximalist agenda) rather than to provide new exceptions or users' rights.

In the trade treaty context, issues of the nature and scope of protection as well as the extent and scope of users' rights in copyright law are typically addressed from the perspective of rights holders in content industries. Yet enhanced protection will have spillover effects into data-dependent industries, which may not be necessary or desirable. For example, while content industries might consider increased protection for technological protection measures — or "digital locks" — to be essential in the fight against unauthorized copying, the same protection, when applied to compilations of data, could have the effect of overriding the basic copyright principle that facts are in the public domain.

There is a second dimension to data and IP that bears mention. Since sharing, interoperability and securing privacy when required are at the core of designing a national data strategy, technical standards matter. A standard bestows near-monopoly power (and accompanying rents) to the producer of the standard. But Canada's performance in generating IP is dismal. Here is the stark reality: for the past 15 years, Canada's business spending on research and development (R&D) ratio to GDP has been in constant decline. Canada's business R&D, now at 0.88 percent of GDP, is so low that it is half of the Organisation for Economic Co-operation and Development average of 1.64 percent, and less than one third of the world leaders.[9] Worse, many of Canada's inventions, even in strategic areas where we have been the pioneers, such as AI, are now embedded in IP owned by foreign organizations. To put it in perspective, as of July 2017, machine learning-patented application in the last 10 years by Microsoft was 1,030, IBM was 580, Intellectual Ventures — the world's most-feared patent-enforcing entity (colloquially known as patent trolls) was 50 and the total from all Canadian firms, research institutions and individuals put together was 48 (Hinton quoted in Castaldo 2017).

It is important, therefore, that Canada ensure the following:

- The price we pay to use standard-essential IP should be as low as possible.
- The access granted to Canadian companies and innovators to use and improve upon IP should be as wide as possible.
- Intellectual property right (IPR) and technology-standard fluency is a necessary skill for the future. All Canadians should be IPR savvy. China could serve as a model to follow in these efforts.
- Since Canada's stock of research on which high-quality global IPRs are based, is significantly higher than our stock of such IPRs, we should aggressively ensure — through prior research efforts, inclusion in technology standards and expanding efforts to translate Canadian research into internationally recognized IPRs — that our very own inventions are not used to limit the freedom to operate of Canadian companies and entrepreneurs.
- Since the only valuable IPR is an internationally recognized one, Canadian IPR-granting and IPR-regulating organizations should focus more attention on enhancing and increasing the level of Canadian-owned IPR that is globally filled and/or registered by all means possible.

9 See www.oecd.org/sti/outlook/e-outlook/sticountryprofiles/canada.htm.

- International standards are not a policy tool limited to high-technology industries. As the Chinese government, and multiple American corporations, amply demonstrated in the last decade, technology standards are important in all areas, not least in agriculture. Canada should develop a cohesive strategy that sees international standards setting as part and parcel of its innovation and growth policy in all sectors and regions.
- The involvement of both Canadians and Canadian technologies in international technology standards, their governance and their setting should be maximized. The aim here would be a dual-pronged strategy to increase the levels of Canadian IPRs and research embedded into international standards, as well as influencing the norms governing the setting and the payment on patent-essential IPRs to maximize the access and minimize the cost for Canadian firms and individuals.

On the last point, there is some good news: Canadian organizations under the leadership of the Standards Council of Canada have been increasing their involvement in this important arena. However, that alone is not enough to back the IP trade imbalance trend. We need to take a page from the Chinese playbook: if Canada wants to enjoy sustained prosperity in the age of the knowledge economy, it is not enough to have a side seat at the table, we need to take leadership positions in global fora.

International Economic Agreements, including NAFTA Renegotiations

Treaties last for decades, so they must be viewed as inherently rigid and permanent. It is critical to thoroughly understand the kinds of effects the proposed "data elements" of these "twenty-first-century trade agreements" will have on Canada's economy, privacy, sovereignty and democracy. Given that the role of data in our lives is rapidly evolving, policy flexibility is critical because we do not currently know what we will need even five or 10 years from now. Since Canada has few large technology companies, negotiators need to be especially vigilant to guard against foreign multinational corporations' push for a "maximalist agenda."

The recent passage of the Defend Trade Secrets Act of 2016 in the United States reflects its growing interest in creating a more robust national framework for the protection of trade secrets, and this will likely be a US priority to integrate into NAFTA.

In NAFTA, Canada should ensure the inclusion of language on "maintaining balance across all IP rights, the legitimate interests of users, promoting access to and preserving the public domain, ensuring that IP rights do not create barriers to legitimate trade and facilitating access to affordable medicines" (Geist 2016a). Similar language was raised during the Trans-Pacific Partnership negotiations and it belongs in NAFTA.

The NAFTA IP chapter should also "address the abuse of intellectual property rights that may inhibit companies from innovating or discourage Canadians from taking advantage of the digital market" (Geist 2016b). The benefits of an anti-IP abuse law could be used to touch on patents, trademarks and copyright.

One of the chief concerns with past trade negotiations is the expectation that the United States requires other countries to mirror its IP laws, even if those laws extend far beyond international law requirements. The Canadian approach should be to require NAFTA parties to meet international law, but to retain the full flexibility found within those laws.

Works Cited

Bohaker, Heidi, Lisa Austin, Andrew Clement and Stephanie Perrin. (2015). *Seeing through the Cloud: National Jurisdiction and Location of Data, Servers, and Networks Still Matter in a Digitally Interconnected World*. www.ryerson.ca/content/dam/politics/docs/internal/SeeingThroughTheCloud-BohakerAustinClementPerrin-150915.pdf.

Cadwalladr, Carole. 2017. "The great British Brexit robbery: how our democracy was hijacked." *The Guardian*, May. 7. www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy.

Castaldo, Joe. 2017. "Why does Canada give away its best ideas in AI?" *Maclean's*, April 13. www.macleans.ca/politics/ottawa/why-does-canada-give-away-its-best-ideas-in-ai/amp/.

Dwoskin, Elizabeth, Craig Timberg and Adam Entous. 2017. "Russians took a page from corporate America by using Facebook tool to ID and influence voters." *The Washington Post*, October 2. www.washingtonpost.com/business/economy/russians-took-a-page-from-corporate-america-by-using-facebook-tool-to-id-and-influence-voters/2017/10/02/681e40d8-a7c5-11e7-850e-2bdd1236be5d_story.html?utm_term=.6a81c9a4b3df.

Farand, Chloe. 2017. "House Republicans finalise controversial internet privacy law letting US companies sell browsing history." *The Independent*, March 30. www.independent.co.uk/news/world/americas/us-politics/republicans-internet-privacy-law-passes-house-user-data-a7657866.html.

Geist, Michael. 2016a. "The Trouble with the TPP, Day 1: U.S. Blocks Balancing Objectives." January 4. www.michaelgeist.ca/2016/01/the-trouble-with-the-tpp-day-1-u-s-blocks-balancing-objectives/.

———. 2016b. "Ottawa needs to combat abuse of intellectual property rights: Giest." *The Toronto Star*, August 15. www.thestar.com/business/2016/08/15/ottawa-needs-to-combat-abuse-of-intellectual-property-rights-geist.html.

Johnson, Erica. 2017. "TD insider says bank doesn't want you to know it's outsourcing work overseas." CBC News, October 27. www.cbc.ca/news/business/td-bank-outsourcing-fraud-india-1.4362324.

McBride, Kurtis. 2015. "The big tech firms are wrong." CityMetric, November 18. www.citymetric.com/horizons/big-tech-firms-are-wrong-we-must-build-smart-cities-ground-1590.

Owen, Taylor. 2017. "Is Facebook a threat to democracy?" *The Globe and Mail*, October 19. https://beta.theglobeandmail.com/opinion/is-facebook-a-threat-to-democracy/article36661905/.

The Economist. 2017. "Regulating the internet giants: The world's most valuable resource is no longer oil, but data." *The Economist*, May 6. www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource.

The White House. 2017. "Executive Order: Enhancing Public Safety in the Interior of the United States." January 25.

Townsend, Anthony M. 2014. *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia*. New York, NY: W. W. Norton & Company.

About CIGI

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

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

À propos du CIGI

Au Centre pour l'innovation dans la gouvernance internationale (CIGI), nous formons un groupe de réflexion indépendant et non partisan qui formule des points de vue objectifs dont la portée est notamment mondiale. Nos recherches, nos avis et l'opinion publique ont des effets réels sur le monde d'aujourd'hui en apportant autant de la clarté qu'une réflexion novatrice dans l'élaboration des politiques à l'échelle internationale. En raison des travaux accomplis en collaboration et en partenariat avec des pairs et des spécialistes interdisciplinaires des plus compétents, nous sommes devenus une référence grâce à l'influence de nos recherches et à la fiabilité de nos analyses.

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

Centre for International
Governance Innovation

67 Erb Street West
Waterloo, ON, Canada N2L 6C2
www.cigionline.org