Centre for International
Governance Innovation

Policy Brief No. 145 — December 2018

Canada Needs Standards to Support Big Data Analytics

Michel Girard

Key Points

- Standards can bring clarity on definitions, systems architecture, data ownership, grading, pooling, storage, disposal and set the bar regarding privacy and aggregation requirements.
- They are a necessary pre-condition for interoperability and commoditization to occur throughout value chains and across sectors.
- Canada has virtually no institutional capacity to develop standards in the information and communication technology (ICT) sector.
- This policy brief proposes the creation of a standards collaborative that would be entrusted with the development of a standards roadmap to support big data analytics.

Introduction

If big data is indeed "the new oil," new standards, specifications and conformity assessment programs are required, along with changes to administrative/contract law, new legislative frameworks and international agreements. Standards can provide much-needed guidance on the handling of both personal data and data generated by sensors, mechanical devices, images and live camera feeds. Clarity and consistency are required on definitions and ontology, on measurements and on metrics. Verifiable and enforceable rules regarding consent, data ownership, aggregation, protection, storage and disposal are necessary to establish a level playing field.

Lack of standards is a barrier to the growth of the sector and until they are in place, most owners and custodians will be unwilling to share data, supply will remain elusive, and opportunities for discoveries and growth will be missed.

While Canadian organizations are sitting on the sidelines, tech giants with free access to data generated through their platforms are taking a decisive lead in deploying new products and algorithms. For decades, Canadian ICT companies have been compelled to adapt to rules set by tech giants through US-led private sector consortia and open-source platforms. The same could happen for big data. Until the creation of the CIO Strategy Council Standards Consortium earlier this year, Canada had no institutional capacity to develop national positions and standards in support of this important sector of the Canadian economy.

About the Author

Michel Girard is a senior fellow with CIGI's Global Economy Program. Michel's work at CIGI relates to standards for big data and artificial intelligence (AI). His research strives to drive a dialogue on what standards are and why they matter in these emerging sectors of the economy. He highlights issues that should be examined in the design of new technical standards governing big data and AI in order to spur innovation while also respecting privacy, security and ethical considerations. He will offer policy recommendations to facilitate the use of big data and AI standards and their incorporation into regulatory and procurement frameworks.

In addition to his work at CIGI, Michel provides standardization advice to help innovative companies in their efforts to access international markets. He contributes to the CIO Strategy Council's standardization activities and was recently appointed to the International Electro Technical Commission's Market Strategy Board, which develops and maintains more than 10,000 international standards for electrical and electronic technologies.

Michel has 22 years of experience as an executive in the public and not-for-profit sectors. Prior to joining CIGI, Michel was vice president, strategy at the Standards Council of Canada where he worked from 2009 to 2018. Previously, he was director of the Ottawa office at the Canadian Standards Association, director of international affairs at Environment Canada, corporate secretary at Agriculture Canada, and acting director of education and compliance at the Canadian Environmental Assessment Agency. He holds a Ph.D. and a master's degree in history from the University of Ottawa.

Moreover, there is almost no experience in Canada in managing open collaborative standards development platforms.[1] Open platforms are essential for a productive dialogue to take place between software engineers and other stakeholder groups interested in making big data analytics work.

Canada should take the lead and develop foundational standards to allow small and large organizations to thrive. A big data analytics collaborative should be created to develop a standards roadmap that would inform Canadians on standardization activities currently under way, identify gaps and make recommendations for action nationally, regionally and internationally. By acting now and by using the right tools, Canada can design a framework that will allow innovative companies to compete while respecting applicable laws and regulations.

What Standards Are and Why They Matter

Although not visible to the average consumer, standards and conformity assessment activities keep the economy running. They cover everything from the size of the simplest screw thread to the most complex information technology network. Standards provide a level playing field for industry and help build trust between participants in supply chains. They serve as a "handshake" between various components of systems and allow for interoperability. Standards also play a pivotal role in protecting the health and safety of consumers in a wide range of sectors, including food and consumer products, infrastructure and the workplace.

Standards set out requirements, specifications, guidelines or characteristics that can be consistently applied to ensure that products, materials, processes and services perform as intended — qualitatively, safely and efficiently. They are drafted in a way that allows another party to test and certify that a product, process or system meets the requirements

1 The Canadian Standards Association recently launched an open collaboration standards platform that could be used in new sectors such as big data analytics. It would allow all stakeholder groups to engage in formalized standards development activities with software engineers for both normative documents and algorithms. The CIO Strategy Council standards consortium is also using an online document development platform called Central Collab.

of a specific standard. Put simply, they make things work, help innovations spread, and facilitate efficient trade among provinces, countries, economic regions and the international community of nations.

Standards are generally developed through a formalized rule-making process involving engineers and other technical experts, regulators and consumer interests. The process aims at balancing competing interests in order to offer a technical solution that is broadly accepted and shares the benefits of technological compatibility as widely as possible.

Many standards bodies were created at the beginning of the twentieth century to support industrialization. After World War II, new international organizations such as the International Organization for Standardization were established as trade liberalization discussions were gaining traction. Today, thousands of standards development organizations (SDOs) are managing over one million national standards and more than 330,000 international standards.

Standards and Regulations

As the vast majority of technical standards are "voluntary" in nature, there is no formal obligation to comply with them. Market forces dictate their use through supply chain contracts or procurement requirements. Framed properly with appropriate third-party certification programs, standards can indeed be used as an alternative to regulations.

Although practices vary between jurisdictions, developed countries do reference a large number of voluntary standards in regulations. For example, more than 7,100 technical standards and codes are referenced in Canadian regulations covering sectors such as occupational health and safety, construction and infrastructure, energy efficiency, the environment, consumer products, electrical, oil and gas, elevators, pressure vessels, medical devices and organic foods. When a standard is referenced in a regulation, it becomes mandatory.

Standardization in the ICT Sector

In its infancy, the ICT sector followed the same path as other industries and relied on the traditional standards development model. However, with digitization in the 1970s, new approaches were needed to quickly set a large number of new standards and specifications to achieve interoperability (Updegrove 2007). Starting in the 1980s, standards consortia organizations began to appear. Approximately 60 percent of all standards and specifications covering the ICT sector were created by consortia, including well-recognized interoperability standards such as USB drives, DVDs, the Blu-ray optical disc format, HTML, UHD, XML, MIDI and PCI Express (Biddle et al. 2012).

The entire edifice of digitization is based on software development and coding. As this new sector appeared, so did new approaches to draft, test and ensure new products' interoperability, from software to code language and apps. Software developers migrated from consortia to open-source platforms such as GitHub, where software can be designed and tested through the help of a community of 24 million developers.

Open Collaborative Platforms for Standards Setting

The need to establish stronger connections between software engineers and other communities of practice has been well documented. Software engineers are looking for ways to better understand users' needs and to integrate a broader range of considerations such as health, safety and security in software.[2] However, traditional consortia and open-source development platforms are not designed to solicit broad public participation (Leveson 2011). New standards consortia such as the CIO Strategy Council are therefore critical for a productive dialogue to take place.

The Need for "Foundational" Standards

Five factors have been identified in favour of foundational standards for big data analytics:

- Innovation is outpacing legal and regulatory frameworks and the ability of regulators to respond to new issues associated with the deployment of disruptive technologies.
- New laws and regulations are required. Governments are developing approaches to frame new issues on their own, but fundamental

2 In a recent feature article in *The Atlantic*, James Somers (2017) reports on concerns from engineers who are seeing "critical systems that were once controlled mechanically, or by people, are coming to depend on code."

principles are not harmonized around the world, leaving both regulators and industry unsure of how to enforce or comply. Inconsistencies in approaches are adding costs for implementation and lack of compliance due to conflicting requirements.

- Big data analytics will become embedded in all industries, including traditional market players. While in the past each sector built a standardization framework in silos, market participants now employ legions of ICT engineers who will work on big data. Foundational documents can underpin innovations in all market segments and allow for interoperability.
- Geopolitical dynamics of increased nationalism are weakening a number of international organizations aimed at supporting globalization through treaties and binding agreements. The standards development community is one of the few stable institutions that provides an international trust mechanism that can balance essential sovereignty concerns with global trade.
- On the other hand, formal, international standard organizations are no longer the place where most emerging, software-based technical interoperability standardization work takes place. Collaborative development methodologies (i.e., open-source development and informal group projects) have become the preferred method for software-based interoperability development. However, these approaches generally do not satisfy regulators' need to adhere to more formal international requirements regarding government use of "international standards" developed in the private sector.

Some of the core issues raised by big data analytics have been tackled by international SDOs to address other ICT applications and could be adapted to this new sector.[3] However, the following topics have not been tackled and would benefit from the development and use of foundational standards:

Information Architecture

Millions of discrete data sets could be used as input for big data analytics. Standards can help structure and categorize shared information environments and data sets, including organizing and labelling categories of data sets to support usability, findability and traceability. Big data analytics will involve complex value chains. Just as with traditional supply chains for tangible products, each segment will have specific roles and responsibilities, which will have to be described and categorized. In addition, data will go through a life cycle from creation to disposal, which will also have to be described and categorized.

Ontology, Semantics, Definitions and Terminology

When industrial sectors were mostly vertical in nature, standards were developed in silos. As a result, a multiplicity of domain-specific semantics, including product terminology, classification and properties were created. With digitization, information is being generated and exchanged across sectors. This leads to a demand for universal semantics, which should follow a common ontological foundation. It is a prerequisite for interoperability.[4]

Data Owners, Custodians and Controllers

Currently, the majority of data owners, custodians and controllers are not willing or legally allowed to share or sell data to third parties — even with "no regrets" caveats. There are concerns as to whether the data they intend to share:

- should be described as "clean" and devoid of errors or biases that could lead to errors, incidents or accidents downstream and expose them to potential legal actions;
- should be described as "raw" and devoid of manipulation or filtering;

3 The IEEE (Institute of Electrical and Electronics Engineers) P7000 series of standards include: IEEE P7000™ – Model Process for Addressing Ethical Concerns during System Design; IEEE P7001™ – Transparency of Autonomous Systems; IEEE P7002™ – Data Privacy Process; IEEE P7003™ – Algorithmic Bias Considerations; IEEE P7004™ – Standard on Child and Student Data Governance; IEEE P7005™ – Standard on Employer Data Governance; IEEE P7006™ – Standard on Personal Data AI; IEEE P7007™ – Ontological Standard for Ethically driven Robotics and Automation Systems; IEEE P7008™ – Standard for Ethically Driven Nudging for Robotic, Intelligent and Autonomous Systems; IEEE P7009™ – Standard for Fail-Safe Design of Autonomous and Semi-Autonomous Systems; IEEE P7010™ – Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems.

4 The challenge will be to determine whether sectoral semantics definitions and terminologies can be used as sources for higher-level big data analytics vocabulary, or whether an entirely new "language" will have to be developed. For example, the IEEE recently published a glossary for discussion of ethics of autonomous and intelligent systems. The glossary outlines and contrasts variations for each term from five different disciplines (ordinary language, computational disciplines, engineering, government, policy and social sciences, and ethics and philosophy) and significant differences have been highlighted. See Jordan, Day and Ingram (2017).

- will be used downstream in an appropriate fashion;
- may result in unintended consequences;
- may provoke the accidental release of sensitive data; or
- could break any laws, including privacy;

Foundational standards would provide guidance and objective tests to demonstrate compliance and limitations on liability for misuse downstream.[5]

Self-declaration for Data Use

A foundational standard that provides guidance on how to make data available is needed. Standardized self-declaration forms for data use would allow for discoverability, accessibility, known quality and consistency of data. As trust in data is an important consideration for users, it would establish parameters for grading the trustworthiness of the source and the data itself (Saenger et al. 2014). It would also allow for the development of data quality rating/grading systems and provide protocols for data sharing. A standardization self-declaration statement would also help streamline internal approval processes to make data available to third parties.

Data Pools, Trusts, Marts and Warehouses

Large companies with multiple offices have already begun to pool data sets into virtual data marts or warehouses. But there are few known instances yet where data from different organizations is being pooled and used. Once data owners, custodians and controllers from different organizations begin to share data, data marketplaces will be created. Data will be pooled and aggregated. Data brokers will buy and sell data. A foundational standard on the features and requirements of data pools, trusts, marts and warehouses will be needed.

Personal Data

There is no consistency between jurisdictions regarding what constitutes personal information and, as a result, as to what constitutes personal data. In order to avoid the non-authorized release of personal information through multiple data sources, foundational standards are needed for data taxonomy, individual consent and control, and data aggregation/disaggregation (Australian Computer Society 2017).

Other foundational standards may be needed to manage risk, provide guidance on linked data sets storage and establish protocols to alert individuals in the event of a breach or re-identification. Technical specifications for individual products or platforms, as well as standards for the testing and certification of data sets, products and algorithms, will also be required.

Standards Roadmap for Big Data Analytics

Canada would gain by developing a standardization roadmap for big data analytics. Roadmaps are routinely developed to support standardization activities in emerging sectors. It would inform Canadians from key sectors of the economy on standardization activities currently under way, identify gaps and make recommendations for action nationally, regionally and internationally.

As a first step, a standardization collaborative should be created — a cross-sector coordinating body whose objective would be to accelerate the development of foundational standards and specifications consistent with stakeholder needs. The collaborative may involve more than 100 organizations, including the Standards Council of Canada; the standards development organizations/consortia (national, regional and international) with a stake in big data analytics; innovative companies and academics; regulators; and representatives from key sectors of the economy with an interest in the issue and civil society. Through the collaborative, participants would:

- test the merits of voluntary "foundational" standards supporting big data analytics that would apply across value chains, sectors and nations;
- gauge interest for the use of online, open and collaborative platforms that adhere to the World Trade Organization's six principles for standards development in order to adapt to the realities of the market and the practices of the ICT sector;
- determine whether regulators would support the development of voluntary foundational standards with the view to consider adopting them in regulatory frameworks in the future; and

5 Data.gov, the home of the US government's open data portal, provides access to more than 302,000 data sets. See www.data.gov.

→ consult with relevant bodies regarding the feasibility of adopting foundational standards as international standards and developing new international conformity assessment and accreditation programs in this area.

The collaborative would work on a standards roadmap outlining the national, regional and international standards landscape for big data analytics, including standards already published and those under development. It would assess gaps and make recommendations for priority areas where additional standardization and/or pre-standardization research and development are needed. The roadmap would allow governments and industry to better ascertain the investments required to develop a sizable corpus of standards to support big data analytics in each of Canada's key economic sectors. The roadmap would be updated periodically to assess progress and identify emerging issues that require standardization. Priority standards development activities could be identified and acted upon as the document is drafted.[6]

Canada is falling behind other developed economies in designing standards strategies for big data analytics. Many developed economies are forging ahead. In Australia, government and industry are working through the Australia Computer Society on standards for data privacy and artificial intelligence (AI) (Australian Computer Society 2017). China has already published a number of national standards covering data privacy, cross-border data transfer, personal data, big data security, data protection and some AI applications (Sacks and Li 2018). The newly created UK Centre for Data Ethics and Innovation will support standards that guide ethical and innovative uses of data and AI.[7] And the European Union has commissioned the European Technical Standards Institute to develop standards to facilitate compliance to the General Data Protection Regulation.

Canada is uniquely positioned to take the lead on a forward-looking international standards agenda and help achieve sound data governance. But it needs to create the right framework nationally to make it happen.

6 See, for example, America Makes and ANSI Standardization Collaborative (2018).

7 See www.gov.uk/government/consultations/consultation-on-the-centre-for-data-ethics-and-innovation/centre-for-data-ethics-and-innovation-consultation.

Works Cited

America Makes and ANSI Standardization Collaborative. 2018. "Standardization Roadmap for Additive Manufacturing." Version 2. June. Washington, DC.

Australian Computer Society. 2017. "Data Sharing Frameworks." Technical White Paper. Sydney, Australia: Australian Computer Society. September.

Biddle, Brad, Frank X. Curci, Timothy F. Haslach, Gary E. Marchant, Andrew Askland and Lyn Gaudet. 2012. "The Expanding Role and Importance of Standards in the Information and Communications Technology Industry." *Jurimetrics* 52 (2): 177–208. www.jstor.org/stable/23239825?seq=1#page_scan_tab_contents.

Jordan, Sara, Rosalie Day and L. Maria Ingram. 2017. "A Glossary for Discussion of Ethics of Autonomous and Intelligent Systems." Version 1. IEEE, October.

Leveson, Nancy. 2011. *Engineering a Safer World: Systems Thinking Applied to Safety*. Cambridge, MA: MIT Press.

Sacks, Samm and Manyi Kathy Li. 2018. "How Chinese Cybersecurity Standards Impact Doing Business in China." CSIS Brief. August. Washington, DC: Center for Strategic and International Studies.

Saenger, Johannes, Christian Richthammer, Sabri Hassan and Gunther Pernul. 2014. "Trust and Big Data: A Roadmap for Research." 1st Workshop on Security in Highly Connected IT Systems (SHCIS), Munich, Germany, September.

Somers, James. 2017. "The Coming Software Apocalypse." *The Atlantic*, September 26. www.theatlantic.com/technology/archive/2017/09/saving-the-world-from-code/540393/.

Updegrove, Andrew. 2007. "ICT Standards Setting Today: A System under Stress." *First Monday* 12 (6).

About the Global Economy Program

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

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

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

About CIGI

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

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

À propos du CIGI

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

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

CIGI Masthead

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

Centre for International Governance Innovation

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