Policy Brief No. 155 - September 2019

Standards for the Digital Economy: Creating an Architecture for Data Collection, Access and Analytics

Michel Girard

Key Points

- → Canadian industry and thought leaders view digitization as a way to enhance the competitiveness of the economy; digitization can also improve the delivery of services such as health care.
- → In order to achieve this vision, new data value chains are needed.
- → Data value chains would allow participants in existing supply chains to share data, gain new insights, solve problems and become more efficient.
- → Standards are required to clarify the roles and responsibilities of participants in data value chains regarding data collection and grading, data access and sharing, as well as data analytics and solutions.
- → Standards are also necessary to achieve interoperability and set appropriate benchmarks regarding data governance both necessary preconditions for data sharing between organizations.

Introduction

Canada's annual GDP growth, like that of other Group of Seven countries, is not keeping up with other regions of the world. To remain competitive, the economy needs a jolt. According to the recently published overview report from Canada's Economic Strategy Tables, digitization could do the trick. Industry and thought leaders from many sectors of the economy have argued for the creation of new data value chains, which would allow participants in existing supply chains to work together and solve problems through big data analytics. By collecting, sharing and analyzing data from a multiplicity of sources, it would be possible to gain new insights about market trends, remove bottlenecks, streamline logistics, gain efficiencies, improve quality and enhance the competitiveness of key sectors of the economy. The right standards could create a level playing field and also help Canada's fledging digital industry sector to scale up.

This policy brief outlines the key recommendations from Canada's Economic Strategy Tables and other think tanks pertaining to digitization and data interoperability standards. They confirm a growing demand for the creation of data value chains. It then proposes a common approach to facilitate the standardization of data collection, data access and data analytics, which could be implemented in specific sectors of the economy. Finally, it outlines some of the key data governance themes that need to be addressed to facilitate data sharing between organizations.

About the Author

Michel Girard is a CIGI senior fellow. 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.

The Case for Data Value Chains

In 2017, the Government of Canada launched its Innovation and Skills Plan as a first step to position the country as a global leader in innovation. With this plan, six Economic Strategy Tables were created to foster innovation in natural resources, manufacturing, agri-food, health and biosciences, clean tech and digital industries. More than 80 industry and thought leaders were asked to describe key challenges facing their respective sectors, and to identify opportunities for growth. The final reports, tabled in the fall of 2018, present similar themes across industries, from regulatory barriers and red tape and limited access to capital and skilled labour to the necessity of expanding export markets.

The reports also focus on the urgency to digitize operations and supply chains in order to remain globally competitive. The overview report, which consolidates recommendations from the individual reports, declares: "All economic sectors must be digital sectors. Bold adoption of digital platform technologies will enable us to leapfrog other countries" (Canada's Economic Strategy Tables 2018a). Indeed, as outlined below, a strong case is being made to develop sector-specific data strategies and to create new data value chains.

Although natural resources (including forestry, energy and mines) are expected to remain central to Canada's portfolio of exports, the "Resources of the Future" report calls for the development of a Canadian data strategy for the natural resources sector to successfully integrate digitization and the Internet of Things (IoT) into supply chains. It argues that digital adoption in the United States was a significant driver behind the growth of lower-cost US oil that disrupted oil markets in recent years. It recommends the installation of advanced digital sensors on oil fields and extraction equipment to boost effectiveness, lower costs and improve safety. Private sector data sharing, data pooling and artificial intelligence (AI) are also proposed to enhance the competitiveness of the sector.

The "Agri-food" report points to "huge opportunities to supply the growing global demand for protein" as the world's global population is expected to reach 10 billion people by 2050 (Canada's Economic Strategy Tables 2018b, 2). But the sector needs digitization to remain globally efficient. Regarding big data analytics, it notes that "Agrifood businesses are adopting digital technologies that collect large amounts of data. Data is being collected but stored in different formats and different platforms. This lack of interoperability inhibits the use of shared open-data platforms that provide important insights and enable new innovations to sprout up" (ibid., 15).

The "Advanced Manufacturing" report notes the continued hollowing-out of the Canadian manufacturing sector and urges for a rapid transition toward digitization. According to the report, the sector faces a stark choice: "it will either adopt technology or die" (Canada's Economic Strategy Tables 2018c, 4). However, "with the right technologies in place — robotics, additive manufacturing and big data analytics — Canadian manufacturers can spur innovation and transform the efficiency of their operations" (ibid., 2). Digital manufacturing will impact virtually every facet of manufacturing: from how products are researched, designed, fabricated, distributed and consumed to how manufacturing supply chains integrate and factory floors operate (ibid., 12).

Digitization is also featured in the "Health and Biosciences" sector report. The report calls for the creation of a national digital health strategy "that will provide a framework for privacy and data security, data governance and data sharing, and increase the information available to patients so they can make decisions about their own health outcomes" (Canada's Economic Strategy Tables 2018d, 10). The report goes on to say that "highperforming interoperable, digital systems are seen as a critical enabler of data-driven advances in health. Artificial intelligence is already being used to create patient-centric treatment plans based on a combination of data analytics and the most recent scientific studies. Digital and data transformation will increasingly play a role in finding active therapies for incurable or difficult-to-cure diseases as well as greater success in targeting specific treatments to individual patients" (ibid., 11).

In the "Digital Industries" report, data is described as "the most lucrative commodity of the new global economy" (Canada's Economic Strategy Tables 2018e, 12). Data analytics and self-teaching algorithms are projected to continue to disrupt every imaginable market; however:

In the absence of clear regulations for data infrastructure and the way data is owned, collected, processed, stored, and used, firms (especially large multinationals) will make their own rules. Canada must act now to backfill decade-old laws and regulations. New policies need to:

- → Balance privacy and data protection with commercial value in international markets;
- → Consider and protect data ethics;
- → Promote equity and equality in the age of algorithms.

Our global success hinges on how Canadian businesses apply data to drive innovation. We lead the world in AI research, but we face intense competition to capitalize on our innovations and ensure long-term prosperity and sovereignty. Our AI leadership needs to focus on commercialization, rather than just research, to overcome this challenge. (ibid., 12-13)

Standardization will help both governments and industry design new data value chains and establish a level playing field that will allow smaller Canadian firms to compete. Standardization will also help modernize data-related regulations; in Canada, thousands of standards are incorporated by reference in federal and provincial regulations as a compliance mechanism.

The Economic Strategy Tables are now starting the implementation phase of their work.¹ They will find a growing number of allies and supporters regarding the creation of data value chains. Collaborative data sharing has recently been identified as a priority by leading national organizations and think tanks. For example, the

¹ The Fall Economic Statement, released in November 2018, included a commitment on the part of the government to continue to engage with the tables in a collaborative manner over the long term.

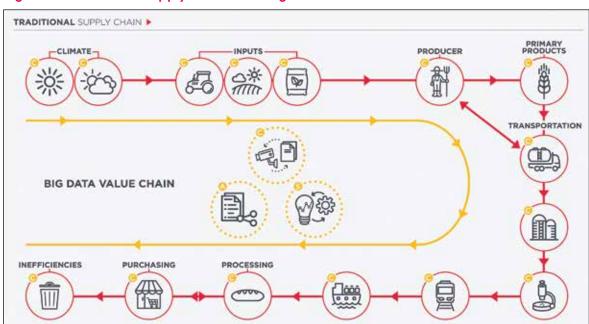


Figure 1: Traditional Supply Chain and Big Data Value Chain

Chartered Professional Accountants (CPA) of Canada has recently noted concerns about the exponential increase in data combined with a lack of standards related to data governance and integrity. It has begun to look at how best to frame assurance regarding data accuracy, quality and value (CPA Canada 2019, 8). A recent report from Deloitte on AI concluded that "certain industries have a clear business case to pool their data to capitalize on far larger datasets than any one company could capture alone without infringing on its customers' rights" (Deloitte, n.d., 21). The recently created Digital Supercluster will consider funding "data commons" projects that involve the aggregation of diverse sources of data to create collaborative platforms on which to build new products.2 Ian McGregor, CEO and chairman of North West Refining, recently proposed the launch of a pilot project for collecting and cataloguing the right data for rapid application of machine learning and AI to Canada's biggest primary industries. The pilot would allow for the creation of the infrastructure and begin populating what could eventually become a large open-source data library for primary industry (McGregor 2018).

Creating New Data Value Chains

If Canada is serious about becoming a digital economy, it needs to develop a new architecture in order to support systematic data collection and grading, data access and sharing, and data analytics among a wide variety of organizations. Data value chains that cut across organizations currently do not exist. The vast majority of data analytics work in Canada is taking place within single organizations. The risks and uncertainties associated with data sharing between organizations, even between divisions or branches in the same organization, are inhibiting data sharing. Foundational standards are needed to bring clarity to intended users across new data value chains, establish common parameters, allow for interoperability, and set verifiable data governance rules to establish and maintain trust between participants and with regulators.

² Five digital superclusters were created by the federal government in 2017 through a five-year, \$800 million investment, in order to foster innovation and create high-paying middle-class jobs.

Ideally, common foundational standards would be developed to frame three distinct subsets of activities starting with data collection and grading upstream; data access, grading, sharing, storage and retention in the middle of the value chain; and data analytics and solutions downstream. Reviews of current practices and trends point to a segmentation of activities, competencies, expertise and accountabilities along these three subsets.³

In order to examine the concept in more detail, let us build on a user case proposed in the "Agrifood" economic strategy report. With the future deployment of fifth-generation (5G) technologies, the sector sees an opportunity to create new data sets by installing a large number of sensors on transportation and handling infrastructure to increase efficiency, adopt smart technologies and automate. The report points to transportation bottlenecks and vulnerabilities across the country that could be addressed through the application of AI. It recommends new standards to "facilitate an open operating system or multiple, interconnected open operating systems so data can be shared among and analyzed by farmers, food processors, distributors, software vendors, equipment manufacturers and data analytics companies" (Canada's Economic Strategy Tables 2018b, 16).

The illustration in Figure 1 attempts to segment a typical agri-food supply chain (represented by the red icons and arrows) into its main components, from climate information and farm inputs upstream to food processing, sale and waste disposal downstream. It then overlays a hypothetical value chain (represented by the yellow icons and arrows) composed of the three subsets of activities proposed earlier: data collection and grading; data access and sharing; and data analytics and solutions.



Data Collection, Organization and Grading – Data Engineers

Activities related to data collection and grading could be distributed across a large number of organizations currently engaged in the supply chain as well as new organizations collecting new data with specialized equipment such as IoT devices. They are represented by this icon and by the letter C affixed across supply chain participants. Upstream, data could come from specialized weather and climate services offered to food producers; a multiplicity of farm inputs providers; or even from food producers themselves as they make decisions about what crop to plant and report on growth conditions and harvesting. Downstream in the supply chain, a series of intermediaries could offer valuable information regarding product transportation and handling of the product and its quality grading before shipment to the final buyer. Invaluable data regarding the processing, distribution, sale, waste, use and disposal of the product could also be generated, offering new insights to participants upstream. In addition to data collection, this segment of the data value chain is best positioned to provide invaluable information about the characteristics of the data collection apparatus and about data set attributes, in order to precisely describe the features of the data collected; apply a grade to the data to make inferences about its quality; ensure it is tagged with the appropriate intellectual property (IP) and copyright for traceability; and get the necessary authorizations for releasing it to users. Finally, participating organizations will want to apply a financial value to the data collected, determine the asking price

See www.kdnuggets.com/2019/06/7-steps-mastering-data-preparationpython.html; Williamson (2019).

for data use, set up structures to administer data exchange activities and report on data transactions. It is anticipated that both private sector and public sector organizations will collect and share data through new data value chains, each with different objectives, capacities and constraints, which must be reflected in future standards. Data engineers would likely be the dominant professional class managing data collection and grading activities on behalf of organizations.



Data Access, Sharing and Retention – Data Controllers

This second segment of the data value chain is needed to make data accessible and usable. It will serve as the interface to connect data sets with organizations performing analytics and generating insights. New data access organizations will be created to manage and track data flows on behalf of all participants making data available. And they will manage data access for AI and machine learning organizations looking to generate insights and new algorithms. Depending on the needs and constraints of participating organizations, the operations of this segment could be decentralized across a supply chain (for example, through data access models based on credentials) or centralized by physically pooling available data into data lakes, commons, trusts, marts, pools, libraries, and so on. In addition to choices about data access modalities, interoperability issues will have to be addressed by data access organizations. Central to interoperability is the choice of an appropriate Application Programming Interface (API) to allow for data transmission, use and tracking. In 2018, there were more than 450 different IoT platforms available in the

global marketplace, but the number could soon reach close to 1,000 different available platforms (McClelland 2018). Open IoT specifications and architecture could be a solution if they prove to be conducive to interoperability (Margossian 2018). Data access organizations will need to manage four core functions: data integration; systems interoperability; data provisioning; and data quality control. Some of the core functions would include managing authentication and data access filters among participants; managing data integration; administering data cleansing and aggregation functions to meet privacy and other regulatory requirements; managing data cloud, residency and retention policies; designing and operating appropriate data dashboards for access and queries; monitoring data flows and transactions and managing smart contracts between participants; enforcing rules regarding data reuse and data transfers, manage connections with other APIs, and report on activities and outcomes. In some cases, data access organizations may also manage the IP and the licensing of algorithms and solutions on behalf of all participants. Data controllers would likely become the main professional class managing data access organizations.



Data Analytics, Solutions and Commercialization – Data Scientists and AI Specialists

This third segment of activities will be undertaken by a number of organizations from supply chain participants, governments, academic and research organizations, small and medium-sized enterprises (SMEs) engaged in AI and machine learning as well as consulting firms providing advice and insights to specific economic sectors. Depending on how the segment is organized, it could operate in a central location under an umbrella organization such as an IoT lab in order to foster collaboration between participants, or could operate in a decentralized way where each organization becomes a member of the data analytics organization and negotiates appropriate access rights to data with the relevant data access organization in order to access data and determine how best to use it. By relying on IoT labs or commercialization incubators as vehicles for generating data insights, supply chain participants would be able to articulate to AI specialists the most urgent problems to solve. They could provide guidance on data availability and quality and test solutions and insights as they get developed. Data analytics organizations would ensure that algorithms and solutions respect applicable regulations and standards and generate trust. Capital and non-financial resources could be made available to small AI firms looking to scale once insights have been generated. IP pools could be created for the benefit of participating organizations. Data analytics organizations could also manage the commercialization of solutions. Data scientists and AI specialists would likely be the dominant professional classes managing data analytics functions in data value supply chains. Depending on the policies of data access organizations regarding data sharing, data brokers, aggregators and traders may also get involved.

Data Governance Themes

In order to be successful, new data value chains will have to address a number of data governance issues that are currently impeding data sharing and exchange between organizations. Standards framing data collection and grading; data access, storage and retention; and data analytics and solutions will need to provide guidance to organizations on the following themes:

- → ownership/IP/copyright;
- → data quality and valuation;
- → interoperability;
- → safe use;

- → trustworthiness;
- → cyber security;
- → data sovereignty and residency;
- → professional credentials and accountability (data engineers/data controllers/data scientists/ data valuation and assurance); and
- → privacy/human rights/digital identity.

Next Steps

Integrating digitization to operations, supply chains, logistics and infrastructure among multiple organizations will require an unprecedented effort toward standardization in order to succeed. Moreover, if Canada wants its fledging digital industry sector to play a meaningful role in delivering solutions, as opposed to foreign-owned tech giants, it needs to take the lead in developing the required standards both nationally and internationally.

As explained in an earlier policy brief (Girard 2018), standards serve as a "handshake" between various components of systems. They allow for interoperability to take place and build trust between participants in supply chains. Their use makes our devices and products work better, for example, by ensuring that the connection between a smartphone and a Wi-Fi network happens anywhere in the world.

In the case of new data value chains, standards are needed to enable connections between various actors that are not taking place now.

As an immediate next step, the participants in Canada's Economic Strategy Tables should be consulted regarding the adequacy of the model for creating data value chains proposed above. There may be an interest on the part of relevant government departments and agencies in launching pilot projects aimed at resolving issues affecting supply chains through data analytics. Canada's digital industries sector could provide advice and support in the creation of these pilot projects.

As an intermediate step, organizations such as CPA Canada may have an interest in providing guidance regarding data assets valuation and the development of new standards to frame data collection and grading; data access, storage and retention; and data analytics and solutions in a way that can be managed by both SMEs and large organizations alike. CPA Canada could also serve as a designated training agency to frame the certification of data valuation and data assurance professionals as organizations develop this new line of business.

Looking forward, the CIO Strategy Council or other accredited standards development organizations could begin to work on foundational standards framing data collection and grading; data access, storage and retention; and data analytics and solutions in order to provide much-needed guidance on how to set up and connect these essential new functions to transform Canada into a digital economy.

Works Cited

- Canada's Economic Strategy Tables. 2018a. The
 Innovation and Competitiveness Imperative:
 Seizing Opportunities for Growth. September.
 www.ic.gc.ca/eic/site/098.nsf/vwapj/
 ISEDC_SeizingOpportunites.pdf/\$file/
 ISEDC SeizingOpportunites.pdf.
- ——. 2018b. "Agri-food." September. www.ic.gc.ca/eic/site/098.nsf/vwapj/ISEDC_ Agri-Food_E.pdf/\$file/ISEDC_Agri-Food_E.pdf.
- ——. 2018c. "Advanced Manufacturing." September. www.ic.gc.ca/eic/site/098.nsf/ vwapj/ISEDC_AdvancedManufacturing. pdf/\$file/ISEDC_AdvancedManufacturing.pdf.
- ——. 2018d. "Health and Biosciences." September. www.ic.gc.ca/eic/site/098. nsf/vwapj/ISEDC_HealthBioscience. pdf/\$file/ISEDC_HealthBioscience.pdf.
- ——. 2018e. "Digital Industries." September. www.ic.gc.ca/eic/site/098.nsf/vwapj/ ISEDC_Digital_Industries.pdf/\$file/ ISEDC_Digital_Industries.pdf.
- CPA Canada. 2019. Foresight: The Way Forward.
 Toronto, ON: CPA. www.cpacanada.ca/fore-sight-report/en/index.html?sc_camp=0634B-51FD23E4B69A83478F09B7FB5D1#page=1.
- Deloitte. n.d. "Canada's AI imperative: Public policy's critical moment." www2.deloitte.com/content/dam/Deloitte/ca/Documents/deloitte-analytics/ca-public-policys-critical-moment-aoda-en.pdf?location=top.
- Girard, Michel. 2018. Canada Needs Standards to Support Big Data Analytics. CIGI Policy Brief No. 145. www.cigionline.org/ publications/canada-needs-standardssupport-big-data-analytics.
- Margossian, Ana. 2018. "IoT and Smart Cities The State of Play Globally." *IoT For All* (blog), November 29. www.iotforall.com/ smart-cities-globally/.

- McClelland, Calum. 2018. "What is an IoT Platform?"

 IoT For All (blog), August 17. www.iotforall.com/
 what-is-an-iot-platform/?utm_
 source=CampaignMonitor&utm_
 medium=email&utm_campaign=IFA_
 Signup_Welcome_Articles?.
- McGregor, Ian. 2018. "Big Data: The Canadian Opportunity." In *Data Governance in the Digital Age*, Waterloo, ON: CIGI. www.cigionline.org/ publications/data-governance-digital-age.
- Williamson, Cory. 2019. "How to Plan Your IIoT Solution for Long-Term Stability." IoT For All, June 20. www.iotforall.com/ how-to-plan-iiot-solution-scalability/.

About Global Economy

Addressing the need for sustainable and balanced economic growth, the global economy is a central area of Centre for International Governance Innovation (CIGI) expertise. The Global Economy initiative examines macroeconomic regulation (such as fiscal, monetary, financial and exchange rate policies), trade policy and productivity and innovation policies, including governance around the digital economy (such as big data and artificial intelligence). We live in an increasingly interdependent world, where rapid change in one nation's economic system and governance policies may affect many nations. CIGI believes improved governance of the global economy can increase prosperity for all humankind.



DATA GOVERNANCE INTERPORT

Available as a free PDF download and digital essay series at www.cigionline.org/data



Introduction and epilogue by Rohinton P. Medhora

ata has been hailed by some as "the new oil," an analogy that captures the excitement and high expectations surrounding the data-driven economy. The success of the world's most valuable companies (Apple, Google, Facebook and Microsoft) is now underpinned by a sophisticated capacity to collect, organize, control and commercialize stores of data and intellectual property. Big data and its application in artificial intelligence, for example, promises to transform the way we live and work — and will generate considerable wealth in the process. But data's transformative nature also raises important questions around how the benefits are shared, privacy, public security, openness and democracy, and the institutions that will govern the data revolution. The recent Cambridge Analytica scandal has exposed the vulnerability of democracies to data strategies deployed on platforms such as Facebook to influence the outcomes of the Brexit referendum and the 2016 US presidential race. Any national 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. The essays in this collection are grouped into five blocks: the rationale of a data strategy; the role of a data strategy for Canadian industries; balancing privacy and commercial values; domestic policy for data governance; and international policy considerations. An epilogue concludes with some key questions to consider around data governance in the digital age.

Essays by:

Susan Ariel Aaaronson Sachin Aggarwal Dan Breznitz Dan Ciuriak Andrew Clement Norman Doidge Michael Geist Blayne Haggart Ariel Katz André Loranger Ian MacGregor Kurtis McBride Brenda McPhail Jonathan Obar

Taylor Owen Teresa Scassa Amanda Sinclair James Tebrake Bianca Wylie

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 initiatives focus on governance of the global economy, global security and politics, and international law in collaboration with a range of strategic partners and have received support from the Government of Canada, the Government of Ontario, as well as founder lim 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 projets de recherche ont trait à la gouvernance dans les domaines suivants : l'économie mondiale, la sécurité et les politiques internationales, et le droit international. Nous comptons sur la collaboration de nombreux partenaires stratégiques et avons reçu le soutien des gouvernements du Canada et de l'Ontario ainsi que du fondateur du CIGI, Jim Balsillie.

Credits

Director, Global Economy Robert Fay Program Manager Heather McNorgan

Publisher Carol Bonnett
Senior Publications Editor Jennifer Goyder
Graphic Designer Melodie Wakefield

Copyright © 2019 by the Centre for International Governance Innovation

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

For publications enquiries, please contact publications@cigionline.org.



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

Printed in Canada on Forest Stewardship Council® certified paper containing 100% post-consumer fiber.

Centre for International Governance Innovation and CIGI are registered trademarks.

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