Centre for International
Governance Innovation

CIGI Papers No. 237 — January 2020

Standards for Digital Cooperation

Michel Girard

Centre for International
Governance Innovation

CIGI Papers No. 237 — January 2020

Standards for Digital Cooperation

Michel Girard

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 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 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

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

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

Table of Contents

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 the Chartered Professional Accountants of Canada's Foresight Initiative on data governance.

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.

About Global Economy

Addressing the need for sustainable and balanced economic growth, the global economy is a central area of 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.

Acronyms and Abbreviations

5G	fifth-generation
AI	artificial intelligence
APIs	Application Programming Interfaces
COGOV	co-governance architecture
CSTD	Commission on Science and Technology for Development
DNS	Domain Name System
DSB	Digital Stability Board
DSTF	Data Standards Task Force
DTP	Data Transfer Project
ETSI	European Technology Standards Institute
FAAMG	Facebook, Amazon, Apple, Microsoft and Google
FAO	Food and Agricultural Organization
FSB	Financial Stability Board
IAB	Internet Architecture Board
ICAO	International Civil Aviation Organization
ICT	information and communications technology
IEC	Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IEEE SA	Institute of Electrical and Electronics Engineers Standards Association
IETF	Internet Engineering Task Force
IGC	International Grand Committee
IGF	Internet Governance Forum
IMO	International Maritime Organization
IoT	Internet of Things
IP	intellectual property
ISO	International Organization for Standardization
ITU	International Telecommunications Union
JTC1	Joint Technical Committee 1
SARPs	standards and recommended practices
SDGs	Sustainable Development Goals
SDOs	standards development organizations
TC	Technical Committee
UNCTAD	United Nations Conference on Trade and Development
W3C	World Wide Web Consortium
WHO	World Health Organization
WMO	World Meteorological Organization
WTO	World Trade Organization

Executive Summary

Global data standards are urgently needed to foster digital cooperation and manage global tech platforms. As no global organization is currently mandated to coordinate the development of data governance and operations standards, this paper proposes the creation of a Data Standards Task Force (DSTF). Precedents exist where standards development work is coordinated by international organizations in sectors of the economy operating across borders, from aviation and maritime shipping to meteorology, food production, public health and the management of the internet. The DSTF would be entrusted with a dual mandate: enabling the development of technical standards to create data value chains and being accountable for the development of data governance standards needed by regulators to properly frame the leading big tech platforms. The ultimate objective of the DSTF would be to create the required architecture for a "single data zone" where data can circulate freely between participating jurisdictions.

Introduction

This paper makes the case for a new mechanism to coordinate the development of global data standards. Data standards are necessary to achieve digital cooperation. We need standards to create data value chains that string together data collection through collaborative platforms in order to generate insights and solve long-standing problems. Organizations of all sizes, whether public, private or not-for-profit, need a suite of data governance standards to manage issues such as data ownership and use, security, residency, privacy and the protection of fundamental rights.[1] With the right data standards and compliance mechanisms in place, consumers could regain trust in global tech platforms, starting with the FAAMG (the five leading big tech platforms composed of Facebook, Amazon, Apple, Microsoft and Google) and support the deployment of ethical artificial intelligence (AI). Finally, credible data standards can be incorporated by reference in national regulations around the world. This will help avoid an unwieldy patchwork of rules and regulations.

No international organization is currently mandated to coordinate the development, maintenance and use of technical standards covering data value chains and policy-oriented standards covering data governance. Standard-setting activities in this field are fragmented between hundreds of organizations. As no data standards registry is maintained, stakeholders and experts alike are struggling to find out what data standards have been published, who is working on new documents and whether there are gaps that need to be addressed. This makes it extremely challenging to set priorities for future standards work. Although tens of thousands of standards are routinely incorporated by reference in regulations covering established sectors of the economy, regulators managing data governance issues have been hard-pressed to come up with workable alternatives.

Precedents exist where standards development work is coordinated by international organizations in many sectors of the economy operating across borders, from aviation (the International Civil Aviation Organization [ICAO]) and maritime shipping (International Maritime Organization [IMO]) to meteorology (World Meteorological Organization [WMO]), food production (Food and Agricultural Organization [FAO]) and public health (World Health Organization [WHO]). These organizations are mandated by nation-states to address interoperability, health, safety and sustainability issues from a global perspective by applying the "one standard, one test" principle. Significant resources are invested by industry and nation-states for the development and maintenance of tens of thousands of codes, standards, guidelines, specifications and best practices, as well as model technical regulations, which make their way into national regulatory frameworks.

Similarly, the internet and the World Wide Web came about following the creation of new standard-setting bodies such as the Internet Engineering Task Force (IETF) and the World Wide Web Consortium (W3C). In a short period of time, these bodies developed technical standards that led to the creation of one physical network layer and one transport network layer that make up the internet. This is the foundation upon which thousands of applications can operate seamlessly today. A similar approach for standard setting is needed

1 Michel Girard (2019b) provides more information on data value chains and on the key topics that need to be addressed through a standards-based data governance framework.

to anchor international digital cooperation and to regulate FAAMG and other global tech platforms.

Looking forward, new open standards development mechanisms will have to engage with a broader range of stakeholders in a nimble and flexible negotiated rule-making process. We need to move toward "evergreen" data standards to keep up with innovation. We also need to reinvent compliance mechanisms and introduce digital certification processes. Future normative documents supporting data value chains, data governance and ethical AI will have to be machine readable. This is necessary to keep up with governments that are now introducing machine-readable legislation.[2] It is also required for the creation of new digital certification and assurance models that combine the use of algorithms and blockchain.

All of this can only be accomplished with a new, stand-alone data standards cooperation mechanism mandated by governments and applied by industry, including platforms.

The Need for International Digital Cooperation

As data flows increase and the cost of data storage continues to drop, the international community has an opportunity to engage in digital cooperation projects to address systemic problems through new insights generated from big data analytics.

In terms of volume, it is estimated that around 90 percent of all the data in human history was created in the past two years (Marr 2018). In its recent *Digital Economy Report 2019*, the United Nations Conference on Trade and Development (UNCTAD) forecasts that global internet protocol traffic, a proxy for data flows, will grow from 46,600 GB per second in 2018 to 150,700 GB per second in 2022 (UNCTAD 2019, 9). The number of Internet of Things (IoT) devices, which will drive a growing proportion of internet traffic, will continue to rise sharply.[3] In 2018, there were already more "things" (8.6 billion) connected to the internet than people (5.7 billion mobile broadband subscriptions). The number of IoT connections is forecast to exceed 22 billion by 2025 (ibid., 7). According to some estimates, the number of IoT devices could exceed 75 billion later in the decade and reach 200 billion when fifth-generation (5G) technologies are fully deployed.[4] With the continued reduction in data collection and storage costs, new opportunities will arise for data analytics using continuously streaming data, as opposed to traditional data stacked in databases, which requires extensive data processing, cleansing and transfer work (Iny 2019). New data value chains — comprising thousands of participating organizations that support data collection, data access and storage mechanisms, and the production of insights from data to address public good issues — can be created (UNCTAD 2019, 2).

Many global digital cooperation opportunities have been framed through the adoption of the UN Sustainable Development Goals (SDGs).[5] There is potential for pooling data from around the world in areas such as health, agriculture and the environment, to enable data scientists and civil society to better understand complex issues and to find new ways to make progress. A few noteworthy examples of international digital cooperation include the following:

- The WHO, through the World Health Assembly, is proposing a global strategy for digital health, which will be considered in 2020. Participants will look at ways to create a global health digital agenda. Health data could be shared through a common platform to foster better patient-centric outcomes. Digital collaboration projects focusing on Alzheimer's and hypertension will be explored.[6]
- The International Center for Tropical Agriculture launched the Platform for Big Data in Agriculture in 2017. It provides ways to share data on agriculture and seeks to transform research and innovation in food security and climate change. Crop modelling data, geospatial data,

2 See Darabi (2018).

3 See www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/.

4 See www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/; www.intel.com/content/dam/www/public/us/en/images/iot/guide-to-iot-infographic.png.

5 See www.un.org/sustainabledevelopment/.

6 See WHO (2019).

Box 1: Addressing UN SDGs through Digital Cooperation

Source: www.un.org/sustainabledevelopment/development-agenda/.

The SDGs are a collection of 17 global goals set by the UN General Assembly in 2015 for the year 2030. In its recently released report, the UN Secretary-General's High-level Panel on Digital Cooperation recommended the creation of digital collaboration platforms to make progress toward the SDGs. Through globally accepted interoperability standards, it would be possible to create data value chains focused on resolving systemic problems. Data collaboratives have already been established to begin to address some of the goals. Looking forward, data-sharing platforms would facilitate the collection of the right data from participating nations, the management of access rights to data to meet regulatory and ethical requirements and provide a space for academics and AI firms to generate new insights. With the right interoperability standards, it would also be possible to exchange data *between* platforms to look at problems from a broader, multi-disciplinary perspective.

livestock data, ontology data and socio-economic data are available through the platform.[7]

- The WMO manages a global platform to access local meteorological data through its OSCAR (Observing Systems Capability Analysis and Review Tool) interface. Local weather station owners and operators register through a detailed application process using the WMO Integrated Global Observing System Metadata Standard, a semantic metadata standard that provides end-users with detailed information on the features of participating weather stations, the equipment they use as well as the type of data collected.[8]
- On the environment front, cheaper sensors generating more data — and better AI algorithms to analyze it — can further improve our understanding of how complex environmental systems interact and the likely impacts of climate change. One recent example involves inexpensive, portable air-quality sensors and a platform to share results through the Web.[9]
- Data collaboratives can also improve economic inclusion. The India Stack gives government agencies and entrepreneurs the technological building blocks to improve service delivery and develop new business models that promote economic inclusion.[10]

As more initiatives are launched, new global interoperability standards for data value chains are required. According to the UN High-level Panel on Digital Cooperation, interoperability standards are needed to connect data to appropriate

7 See https://bigdata.cgiar.org/.

8 See WMO (2017).

9 See Plautz (2018).

10 See https://indiastack.org/about/.

access and storage platforms upstream, and to AI experts downstream. To be successful, data commons will require criteria for establishing relevance to the SDGs, new standards for interoperability, rules on access and safeguards to ensure privacy and security.[11] This can only be accomplished through a new, stand-alone DSTF.

What Standards Are and Why They Matter

As explained in the CIGI paper *Big Data Analytics Need Standards to Thrive*, standards and conformity assessment activities keep the economy running (Girard 2019a). They cover everything from setting the size of the simplest screw thread to managing 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 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 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 (ISO), were established as trade liberalization discussions gained traction.

In addition, a large number of sector-specific agencies were also created under the auspices of the United Nations after World War II. These agencies have been mandated by nation-states to address interoperability, health, safety and sustainability issues. Significant resources are invested by industry and governments to develop and maintain many thousands of codes, standards, guidelines, specifications and best practices, as well as model technical regulations that make their way into national regulatory frameworks. In addition to these international bodies, thousands of standards development organizations (SDOs) manage more than one million national standards and more than 330,000 voluntary international standards.

The Case for Global Data Standards

The call for voluntary global standards to frame global data collaboration comes from a number of areas. Innovation is outpacing legal and regulatory frameworks and regulators' ability to respond to new issues. Although fast-paced changes are occurring in the tangible economy, processes generally exist for mediation to take place between industry and regulators. In the case of data, the pace of innovation is unprecedented, and no international mechanism is in place for that "mediation" to take place between the digital industries sector and regulators.

Governments are responding by developing approaches to frame new issues on their own, but fundamental principles are not harmonized around the world, leaving both regulators and big tech platforms unsure of how to enforce or comply. Inconsistencies in approaches are adding costs for framework implementation and contributing to a lack of compliance because of conflicting requirements.

Big data analytics is not the exclusive domain of big tech platforms but is becoming embedded

11 See Digital Cooperation (2019).

Box 2: International Agencies with Standard Coordination and Setting Responsibilities

There are hundreds of international agencies responsible to draft technical regulations and develop codes, standards and best practices in specific sectors. Below are three examples of organizations with large codes and standards development portfolios. Demand for technical, process and interoperability standards came from regulators and industry. National governments made binding commitments through treaties and conventions. International bodies were created to coordinate standards development activities. Once developed, normative documents are incorporated by reference in laws and regulations.

ICAO: The ICAO was created through the adoption of the Convention on International Civil Aviation. The ICAO Council adopts standards and recommended practices (SARPs) concerning air navigation, air navigation infrastructure, flight inspection, prevention of unlawful interference and facilitation of border-crossing procedures for international civil aviation. It manages more than 12,000 SARPs, which provide the fundamental basis for harmonized global aviation safety and efficiency in the air and on the ground, the worldwide standardization of functional and performance requirements of air navigation facilities and services, and the orderly development of air transport.

IMO: The IMO was created through the adoption of the Convention establishing the Inter-Governmental Maritime Consultative Organization. It is responsible for the safety and security of shipping and the prevention of marine and atmospheric pollution by ships and manages the development and implementation of treaties and conventions, technical regulations, codes and guidelines related to marine safety and security.

Codex Alimentarius Commission: The commission was created by the FAO and the WHO in 1962 to develop food standards, guidelines and codes of practice aimed at protecting consumer health, ensuring fair trade and promoting co-ordination of all food standards work undertaken by intergovernmental and non-governmental organizations. The Codex covers all foods, whether processed, semi-processed or raw. In addition to 223 standards and 53 codes of practice for specific foods, the Codex Alimentarius contains more than 10,700 specifications covering matters such as food hygiene, food additives and pesticide residues, and procedures for assessing the safety of foods derived from modern biotechnology. It also contains guidelines for the management of governmental import and export inspection and certification systems for foods. Global trade of food products in real terms has doubled over the last 20 years, rising to US$1.6 trillion in 2016.

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 information and communications technology (ICT) software engineers and data scientists to work on big data analytics. Foundational documents can underpin new innovations in all market segments and allow for interoperability, not only with big tech platforms but also between other players.

The geopolitical dynamics of increased nationalism are weakening a number of international organizations aimed at supporting globalization through treaties and binding agreements. The international standards development community is one of the few stable institutions providing an international trust mechanism able to balance essential sovereignty concerns with global trade, because it is in the business of developing voluntary normative documents.

If we do not pre-emptively establish normative standards to help society manage the risks accompanying big data, the consequences will almost certainly be unintended and unanticipated harm. The difference between these digital innovations and historical innovations in the tangible goods economy is that the unprecedented rate of progress and innovative possibilities today can outpace sober second thoughts.

No New UN Body in Sight

As indicated above, there are many opportunities for governments and industry to solve complex problems through international digital cooperation. A comprehensive suite of standards is needed to create data value chains and manage complex governance issues. Ideally, a new international agency created through an international convention would set standardization priorities and coordinate standards development activities. However, as outlined below, there is no consensus among nations to create a formal framework to manage data governance through the United Nations. New approaches involving a coalition of willing partners sharing similar values should be considered in order to make progress.

The United Nations began to discuss data governance issues stemming from the deployment of the internet almost 20 years ago in response to a request by the International Telecommunications Union (ITU), the main standard-setting body reporting to the United Nations (UN General Assembly 2002). The ITU convened a first World Summit on the Information Society, which took place in two phases in 2003 and in 2005. The declarations stemming from the summit called for the implementation of a bold plan of action to build the information society and "to put the ICT sector at the service of development."[12]

Over the next decade, UN-affiliated organizations implemented a series of actions to make the internet more accessible in developing countries, for example, by striking partnerships to introduce broadband internet services around the world. Although progress was made on that front, there was also a growing recognition new public policy issues were emerging as a result of this technology in a number of clusters, including security, human rights, legal, economic, developmental and sociocultural.

Commission on Science and Technology for Development

As a result of a UN General Assembly resolution in 2012 taking stock on progress over the past 10 years, the United Nations mandated the Commission on Science and Technology for Development (CSTD) to undertake "a comprehensive mapping of international Internet public policy issues, the mechanisms dealing with these issues and potential gaps in those mechanisms" (CSTD 2015, 2). Over a three-year period, the CSTD developed a spreadsheet that identified 41 issues bundled into seven clusters. It identified 680 governance mechanisms already in place to deal with these issues, including organizations, policy processes and instruments. It noted that the database it created to catalogue those mechanisms "does not attempt to deliver an exhaustive list… given the breadth and constant evolution of the field of Internet Public policy" (ibid., 5).

Technical standards featured prominently in the CSTD report, which describes two layers concerned with the core functionality of the internet: the physical network layer (telecommunications infrastructure standards) and the transport network layer (technical standards such as the Transmission Control Protocol/ Internet Protocol, Domain Name System [DNS]). They set the foundation for the application network layer (content and application standards) where big tech and e-commerce platforms operate (ibid., 8).

The report looked at communications infrastructure standards, technical standards, web standards, internet protocol numbers, the DNS, the root zone, net neutrality, cloud computing, convergence and the IoT. On the one hand, the CSTD report noted that a handful of bodies were created with the specific purpose to develop the technical standards that were essential to achieve interoperability between the physical, transport and application layers, and between regions of the globe to link various communication networks into one seamless web. On the other hand, it noted gaps in the development of these standards. As regulators and non-technical entities did not participate meaningfully in their development, non-technical aspects, such as human rights, competition policy and security may not have been properly incorporated, leaving a governance gap for others to solve (ibid., 11).

As it concluded its inventory work, the CSTD was unable to reach consensus among participants on an appropriate policy response to address significant gaps in data governance. It noted that "the complexity and political sensitivity of the topic did not allow the group to finalize a set of recommendations on fully operationalizing enhanced cooperation" (ibid., 4).

The CSTD report highlights important findings regarding the critical role that technical standard-setting bodies play in setting the rules for a new

12 See www.itu.int/net/wsis/docs/geneva/official/poa.html.

technology to be deployed globally, such as the rise of the internet and the World Wide Web. Early on, a number of independent and, in some cases, competing communication networks were created, which could not connect to one another. Over time, each constituent agreed to be bound by technical standards developed and maintained by a small number of standard-setting bodies that were created for the sole purpose of setting one set of rules to launch one internet. Through these technical standards, one global physical network layer was established, setting the base for one global transport network layer. A loose governance structure was set up through the following standard-setting bodies:

- The Internet Society is the overall coordination body for the Internet. It was created in 1992 at the bequest of the US Department of Defense. Its global headquarters are located in the United States, with an office in Switzerland and regional bureaus covering Latin America and the Caribbean, Africa, Asia, North America and Europe. It operates as a not-for-profit organization and has a global membership base of more than 100,000 organizational and individual members.[13]
- The Internet Society houses the IETF and the Internet Architecture Board (IAB):
  - The IETF provides technical direction for the internet. It develops technical standards through an open and consensus-based process that is often described as rough consensus and running code. In 2014, a search identified 10,263 normative documents produced by the IETF.[14]
  - The IAB oversees the work of the Internet Research Task Force, which also develops technical standards.[15]
- World Wide Web standards are set up by the W3C, which is open to both organizations and individuals.[16]
- Internet Protocol Numbers are managed bythe Internet Assigned Numbers Authority using a standardized protocol.[17]
- The internet DNS is based on IETF standards and recommendations.[18]

It should be noted that these engineering standard-setting bodies are not mandated to address the policy and governance issues that began to surface as the internet grew in popularity and as new applications and global social media and e-commerce platforms were deployed. The governance issues identified by the CSTD have therefore not found a "home" and have not been addressed since the report's publication in 2015.

UN High-level Panel on Digital Cooperation

In 2018, UN Secretary-General António Guterres convened a High-Level Panel on Digital Cooperation to take a fresh look at this issue. Co-chaired by Melinda Gates (Bill & Melinda Gates Foundation) and Jack Ma (Alibaba Group), the panel looked at ways to strengthen cooperation in the digital space among governments, the private sector, civil society, international organizations, academia, the technical community and other relevant stakeholders (Guterres 2019).

The panel's report notes the lack of global standards supporting data value chains and data governance. It acknowledges that standards are needed to create international digital collaboratives in support of the UN SDGs. It states that ad hoc responses could fragment the interconnectedness that defines the digital age and that competing standards and approaches would reduce trust and discourage cooperation. It therefore proposes "upholding open standards and interoperability to facilitate collaboration" as one of the values that should shape the development of global digital cooperation (Digital Cooperation 2019, 12).

The high-level panel examined the need for technical standards to create new value chains and address interoperability issues. The following statements in the report are noted:

- It acknowledged that new technical standards for data value chains are needed in order to pool data from around the world and create data commons to resolve problems in areas such as health, agriculture and the environment.

13 See www.internetsociety.org/internet/who-makes-it-work.

14 See www.internetsociety.org/about-the-ietf/; http://www.arkko.com/tools/allstats/.

15 See www.iab.org/; https://irtf.org/.

16 See www.w3.org/.

17 See www.iana.org/.

18 See www.ietf.org/rfc/rfc1035.txt.

- Data commons require standards for interoperability, rules on access and safeguards to ensure privacy and security.
- There is scope to launch collaborative projects to test the interoperability of data, standards and safeguards across the globe.
- There is an opportunity to take another look at existing digital infrastructure protocols and standards in order to bring more people online.

The report also calls for the development of new data governance standards to address critical gaps:

- It proposes the creation of audits and certification schemes to monitor compliance of AI systems with technical and ethical standards.
- Regarding consumer protection issues, it points to the lack of international standards and effective compliance mechanisms for the exchange of data in order to better manage data flows; standards regarding the interoperability of mobile money systems; as well as standards for managing data consent when children use devices.
- Regarding cyber security, it notes that while many best practices and standards exist, they often address only narrow parts of a vast and diverse universe that ranges from talking toys to industrial control systems.
- It proposes the development of credible data governance standards to rebuild consumer trust in big data analytics given recent privacy breaches.

The report examined what the United Nations could do to enhance digital cooperation. It noted the growing number of mechanisms accountable to generate norms, standards, policies and protocols in the digital space. In 2015, as noted above, the CSTD identified 680 distinct organizations operating around the internet. Four years later, the estimated number had risen to more than a thousand, resulting in an even more fragmented and diffuse standards cooperation and governance landscape (ibid., 6).

The report refrains from recommending the creation of a new, stand-alone body accountable to coordinate the development of suitable data standards for data value chains and data governance. There does not seem to be the appropriate level of support to warrant the creation of a new UN-based organization to coordinate data standards (both technical and governance), even though they are understood to be distinct from internet standards. Rather, the report examined three options (outlined below) and proposes the launch of a bottom-up stakeholder engagement process to design an appropriate global digital cooperation architecture, including governance mechanisms, funding models and modes of operation. It is not clear how the proposed options would address the need for new technical standards to frame data value chains and manage data governance.

Following the release of the report, organizations such as the Geneva-based Digital Watch Observatory took a closer look at the recommendations and organized an international workshop entitled Unpacking the High-Level Panel's report: Contributions from Geneva.[19] A number of actions were proposed to make progress on the key recommendations of the report and to test the international appetite for pursuing one of the three proposed governance mechanisms. Unfortunately, no specific action was put forward on the need to enhance coordination and collaboration on the standards front. Standards and conformity assessment, the foundation upon which international digital cooperation mechanisms must be articulated, appears to have been left for others to settle.

Elusive Standards for Data Value Chains and Governance

Global technical and governance data standards will contribute to lowering the costs of setting up and operating digital collaboration platforms and, through interoperability standards, will facilitate data sharing between users and between platforms. However, without an international standards coordination body focused on developing the right standards, digital cooperation will remain a pipe dream. Thousands of global technical standards were necessary to support the creation of the internet and the World Wide Web. A large number of standards will also be required to create international digital cooperation platforms in all major sectors of the economy and manage a broad range of data governance issues.

19 See https://dig.watch/events/unpacking-high-level-panels-report-contributions-geneva.

Box 3: Three Options Proposed by the UN High-level Panel on Digital Cooperation for the Establishment of a Data Governance Mechanism Using Existing Institutions

Option 1: Internet Governance Forum (IGF) Plus (IGF Plus) would build on the existing IGF, which was established by the World Summit on Information Society in Tunis in 2005. The IGF is a forum for policy dialogue. It represents the main global space convened by the United Nations for addressing internet governance and digital policy issues. It has a small secretariat based in Geneva, Switzerland, and has been designed to be seen as a neutral, non-duplicative and non-binding process to facilitate the exchange of information and best practices and to identify issues and make known its findings, to enhance awareness and build consensus and engagement. It achieves this dialogue through its annual meetings, topic-specific workshops, dynamic coalitions, best practice forums and other engagement mechanisms. A Best Practices Working Forum on IoT, big data and AI was created in 2018 and meets regularly. The high-level panel proposed that the mandate of the IGF could be expanded by adding a Policy Incubator, which would create necessary policies and norms for public discussion and adoption by regulators. Although its focus is currently articulated around the internet in general, its mandate could presumably be broadened to data governance, data value chains and the regulation of the FAAMG platforms. The IGF Trust Fund would be a dedicated fund for the IGF Plus. There is no mention of the need for data standards coordination.

Option 2: Distributed co-governance architecture (COGOV) relies on the self-forming "horizontal" network approach used by the IETF, the Internet Corporation for Assigned Names and Numbers, the W3C, the Regional Internet Registries, the Institute of Electrical and Electronics Engineers (IEEE) and others to host networks to design norms and policies. This proposal would extend a network approach to issues affecting the broader digital economy and society. The COGOV architecture decouples the design of digital norms from their implementation and enforcement. It seeks to rapidly produce shared digital cooperation solutions, including norms, and publish them for stakeholders to consider and potentially adopt. These norms would be voluntary solutions rather than legal instruments. In themselves, the COGOV networks would not have governing authority or enforcement powers. However, the norms could be taken up by government agencies as useful blueprints to establish policies, regulations or laws. The COGOV would aim to establish clear guardrails for digital technologies. It would identify digital governance issues, form digital cooperation networks and support networks through digital cooperation platforms. Once developed, norms would be deployed by nation-states through laws and regulations. Governments would adjudicate resolve disputes and conflicts.

Option 3: The proposed Digital Commons Architecture would aim to synergize efforts by governments, civil society and businesses to ensure that digital technologies promote the UN SDGs and to address risks of social harm. It would comprise multi-stakeholder tracks to create dialogue around emerging issues and communicate use cases and problems to be solved to stakeholders, and an annual meeting to act as a clearing house. Each track could be owned by a lead organization. Light coordination of the tracks, and servicing of the annual meeting where their reports are considered, could be ensured by a small secretariat housed within the United Nations. Setting norms would be coordinated through the annual meetings where the output of the various tracks would be discussed as well as implementation of the governance guidance produced by these tracks through a "soft" review of reports by stakeholders. Once again, this falls short of the creation of new standards bodies for the coordination or development of data standards, such as the IETF and W3C did for the internet.

Standard-setting activities in the ICT sector can only be described as extraordinarily complex, opaque, evolutionary, bottom up and unpredictable. In addition to hundreds of established SDOs, the sector also relies an even larger number of standard consortia and open-source software development platforms. As a result, making sense of standardization activities covering data value chains and data governance around the globe will require a large-scale effort.

There are no registries of standards and conformity assessment programs in place and academic studies on the subject are sparse. In Canada, the Standards Council of Canada recently announced the creation of a standardization collaborative on big data analytics in order to assemble an inventory of available standards and report on standardization activities currently taking place in the sector.[20] This inventory is undertaken manually. Armed with that information, four working groups will identify standardization gaps and propose a standards road map to fill some of the critical gaps identified through the inventory. The inventory work will prove to be labour intensive given the absence of a central registry on data standards and conformity assessment programs.

A cursory review reveals a dozen major international standards bodies and consortia involved in developing standards and specifications related to big data analytics and some of the key data governance issues that need to be addressed, notably ethical dimensions of AI and privacy. However, no comprehensive standards development activities supporting digital cooperation and the creation of data value chains have been uncovered.

SDOs

In 1987, the ISO and the International Electrotechnical Commission (IEC) established the Joint Technical Committee 1 (JTC1) by merging the ISO Technical Committee (TC) 97 (Information Technology) and the IEC TC 83 (Information Technology Equipment). The JTC1 is seen by many as the leading body making progress in coordinating activities for data management, big data and AI. Its purpose is to develop, maintain and promote standards in the fields of information technology and ICT. Since its creation, the JTC1 has published more than 3,200 standards and publicly available specifications covering a wide array of subjects including programming languages, interconnection of information technology equipment, user interfaces, cloud computing, cyber security, data security, big data, data management and interchange and, more recently, the IoT and AI.[21]

The JTC1 manages a substantive proportion of the two organizations' standards catalogue (ISO maintains more than 20,000 standards and the IEC more than 10,000). The JTC1 operates through a matrix of subcommittees, working groups and advisory groups, which are connected to more than 100 liaison bodies. For example, Subcommittee 42 focuses on big data and AI through four working groups:

- Working Group 1: Foundational standards (concepts and terminology)
- Working Group 2: Big data (overview, definitions, reference architecture)
- Working Group 3: Trustworthiness (biases in AI systems, overview, robustness of neural networks)
- Working Group 4: Use cases and applications

The Institute of Electrical and Electronics Engineers Standards Association (IEEE SA) has been active in the ICT sector for decades through a large number of technical standards for electronic products, such as the Ethernet and WiFi, as well as software engineering management. In 2017, the IEEE had more than 1,100 active standards, with over 600 standards under development. Regarding big data analytics, the IEEE launched in 2017 a global consultation and outreach initiative called Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. The IEEE is now spearheading the development of 15 ethical AI standards under its 7000 series ranging from algorithmic bias consideration to automated facial analysis technology with the help of more than 2,000 participants.[22]

The IEEE SA also launched the development of an Ethics Certification Program for Autonomous and Intelligent Systems, which represents the first attempt to design and deploy an international compliance mechanism toward ethical AI standards. If successful, the new program could provide certification for algorithmic bias, accountability and transparency.[23]

20 See www.scc.ca/en/news-events/news/2019/leading-experts-join-scc-effort-transform-data-governance-landscape.

21 See www.iso.org/isoiec-jtc-1.html.

22 See https://ethicsinaction.ieee.org/.

23 See https://standards.ieee.org/industry-connections/ecpais.html.

In 2018, the IEEE led the creation of OCEANIS, the Open Community for Ethics in Autonomous and Intelligent Systems, along with 15 SDOs that joined as founding members and 19 members from the private sector. It is designed to act as a high-level global forum for discussion, debate and collaboration for organizations interested in the development and use of standards to further the development of autonomous and intelligent systems. Its creation could spur greater collaboration and cooperation among standard-setting bodies focusing on algorithms, sensors, big data, ubiquitous networking and technologies.[24]

The ITU, the UN agency accountable for global standards covering telecommunications and ICT is the custodian of the International Telecommunication Regulations treaty. It maintains more than 4,000 normative documents, including standards. The ITU is an active player in the development of data sharing, IoT and smart cities standards. In 2017, the ITU created a focus group on machine learning for future networks, including 5G, in order to create a unified architecture framework. In 2018, it initiated a focus group on AI for health to create standardized benchmarks to evaluate AI algorithms used in health-care applications.[25]

The European Technology Standards Institute (ETSI) produces standards and specifications for ICT-enabled systems and is focusing on issues such as blockchain, AI, augmented reality and autonomous networks standards. ETSI has published more than 45,000 standards and specifications, which are routinely incorporated by reference in European regulations. It has an ambitious work program related to big data analytics.[26]

The IETF is actively engaged in standardization efforts for Application Programming Interfaces (APIs), IoT devices, privacy considerations, cyber security and metadata insertion.

Standards Consortia and Open Source Platforms

The ICT sector needed a myriad of standards and specifications to deploy digital technologies, hardware, software and the internet. This resulted in the creation of many hundreds of standards consortia. Studies have identified more than 400 ICT consortia operating in that space in the late 1990s (Biddle et al. 2012, 179). As policy research on standards is scarce, it is impossible to know how many data service providers, from IoT device manufacturers, API platforms and AI firms, are managing standards and specifications requirements in the big data analytics space.

There is also a wide array of consortia bodies involved in the development of data standards, some focusing on data architecture, others engaged in sector-specific applications. For example, the Third Generation Partnership Project is developing standards underpinning 5G, IoT narrow-band radio technology and streaming.[27] The Trusted Computer Group develops standards for APIs[28] and The Open Group develops standards for architecture frameworks.[29]

Examples of standards consortia focusing on sector-specific applications include the Clinical Data Interchange Standards Consortium, which deals with medical research data linked with health care, to enable information system interoperability and to improve medical research and related areas of health care;[30] Energistics, which focuses on the development, adoption and maintenance of open data exchange standards for the oil and gas exploration and production industry;[31] and SAE International is creating a consortium to develop best practices and standards for storing and sharing data acquired from shared micro-mobility services.[32]

Following the entry into force of the General Data Protection Regulation, a growing number of bodies have been created to develop appropriate standards regarding personal data privacy, portability and consent. Among the most promising projects is the

24 See https://ethicsstandards.org/.

25 See www.itu.int/en/ITU-T/AI/Pages/default.aspx.

26 See www.etsi.org/committee/1640-sai.

27 See www.3gpp.org/news-events/1607-iot.

28 See https://trustedcomputinggroup.org/wp-content/uploads/TSS_FAPI_v0.94_r04_pubrev.pdf.

29 See https://publications.opengroup.org/standards/togaf.

30 See www.cdisc.org/newsletter/issue/third-quarter-2019/letter-president-and-ceo.

31 See www.energistics.org/solutions/#streamline.

32 See www.sae.org/micromobility/.

Data Transfer Project (DTP) from Google, Facebook, Microsoft and Twitter. The DTP started in 2018 and aims to develop technical standards for personal data portability "so that all individuals across the web could easily move their data between online service providers whenever they want."[33] Once these standards are in place, they could also be used to manage direct and automated data transfers between a source and a data access point, which is a necessary pre-condition for digital collaboratives to operate securely. The objective of another important data privacy standards development project called Solid is to decouple data from applications by offering a new architecture for the web. The project is led by Tim Berners-Lee, the inventor of the World Wide Web. It would allow individuals to choose where their data can be used and for what purpose by creating individual Solid PODs.[34] The European Internet Privacy Engineering Network is also looking at standards development for data privacy.[35]

Finally, open-source platforms have become the main conduit to develop applications for big data analytics, from designing new algorithms to building data-sharing platform software. GitHub, the largest open-source software and coding development platform in the world, now boasts 40 million developers working together to host and review code, manage projects and build software. For example, GitHub hosts more than 34,000 public repositories focusing on machine learning, 25,000 devoted to APIs and 900 projects aimed at building collaborative digital platforms. It has become a major player in defining how big data analytics and digital cooperation will be shaped in the future.[36]

As this cursory review shows, standards activities are fragmented among many organizations. Better coordination would help ensure that the right interoperability standards are developed in order to create data value chains and international digital collaboration platforms. Additionally, data governance issues are not addressed in a systematic way. The current standards corpus would not be sufficient for organizations to develop corporate data policies that propose best practices for pervasive issues such as data ownership, intellectual property (IP) and copyright; data tagging and traceability; digital identity management; privacy and the protection of human rights; ethics; data security; and data residency requirements.

Designing a New Approach

Robust global standards and third-party certification programs are essential to launch an inclusive digital economy. Credible and enforceable global standards are needed for consumers and civil society to regain trust in big tech platforms. They are required to create a level playing field, where smaller firms can compete fairly against big tech platforms, and they represent the only available pathway to avoid an unwieldy patchwork of national regulations. With the right global data standards, the benefits of digitization to society can be maximized while the potential harms from global platforms managed by the private sector are minimized. It is unlikely, however, that a new UN data standard-setting body will be formed, as indicated earlier. There is no international consensus for a convention on digital cooperation and the creation of a stand-alone agency. The best-case scenario proposed by the United Nations would be to expand the role of existing agencies or networks. At first glance, neither the proposed IGF Plus, COGOV or the Digital Commons Architecture would be as effective as a dedicated data standards agency.

Under a status quo scenario, where existing SDOs, consortia and open-source software platforms compete to develop industry standards and specifications, the current fragmented approach will continue. In the absence of a concerted effort on the part of regulators to play an active role in setting and enforcing global data governance standards through established organizations, software engineers and data scientists could be expected to continue to migrate away from traditional standard-setting bodies toward GitHub and the like. The sentiment among many software engineers toward traditional SDOs is that a system that retains strong roots in the nineteenth century is ill-suited to meet the demands of the twenty-first century. As a result, a patchwork of national regulations reflecting the divergent interests and value systems will be created. Over time, we may find ourselves with

33 See https://datatransferproject.dev/dtp-overview.pdf.

34 See https://solid.inrupt.com/how-it-works.

35 See https://edps.europa.eu/data-protection/ipen-internet-privacy-engineering-network_en.

36 See https://github.com/marketplace/category/api-management.

three or four mutually exclusive blocs between which little data will be willingly shared.

In the absence of a global commitment to regulate data governance, there is little appetite among stakeholders to create a new data standards coordination body. Peter Cihon (2019), in a technical report entitled *Standards for AI Governance: International Standards to Enable Global Coordination in AI Research and Development*, argued that the JTC1 is in a better position to coordinate ethical AI standards work compared to other, non-World Trade Organization (WTO) sanctioned bodies such as the IEEE. Cihon argued that the challenge will be to attract AI scientists and researchers to participate in standards setting; many feel the field of AI safety is too young to engage in creating norms and compliance mechanisms (ibid.).

It may, however, be possible to create a *regional* data standards coordination body. In a CIGI paper entitled *A Plurilateral "Single Data Area" Is the Solution to Canada's Data Trilemma*, Susan Ariel Aaronson and Patrick Leblond (2019) proposed the creation of an International Data Standards Board. The organization would initially cover Canada, the European Union, the United States and Japan, but could expand to other nation-states. It would be accountable for devising common technical and governance standards. The standards would ensure a high degree of trust in the data-driven economy among individuals, consumers, workers, businesses and governments so that all forms of data could flow freely across borders. The International Data Standards Board would also be responsible for monitoring the single data area. Regular assessments would determine if participating member states are in compliance with the standards. The authors argued that such a body could not operate under the WTO as the issues requiring standardization are not limited to trade. The organization could be set up as a not-for-profit organization and report to a board of directors composed of representatives from participating nation-states and industry (ibid.).

A similar diagnostic and approach were recently proposed by Ian Bremmer (2019) of the Eurasia Group. In his remarks at the GZERO Summit in Tokyo, Japan, Bremmer argued that the market for data and information is no longer global and is breaking in two. We are facing the development of two distinct tech ecosystems: one built by the private sector and loosely regulated by governments under US leadership, and another ecosystem dominated by the state in China. A fault line between the two emerging systems can be seen in data collection, the development of AI, the rollout of 5G, the deployment of IoT devices and defence and retaliation against cyber attacks. Bremmer called for the creation of a "digital WTO" to set future standards for AI, data, privacy, citizens' rights and IP. The United States, Europe, Japan and like-minded countries that believe in online openness and transparency would lead the way in creating such an organization. China would have an economic and security incentive to want to join, "especially if it's the only way Beijing can secure access to developed markets" (ibid.).

Others are looking at creating global standards for data governance mechanisms to focus on core issues. In Europe, a collaborative called A New Governance: Standardization for Data Empowerment is calling for a new international agency to develop global data standards for personal data protection, circulation and portability. Members of the collaborative are concerned that many sectors are developing stand-alone privacy standards. They have noted initiatives in mobility, health care, administration, commerce, finance and insurance, entertainment, energy, telecom, human resources and education. These initiatives are recreating silos, which will be highly detrimental to the main goal of fostering personal data circulation and protection across sectors and boundaries (Privacy Tech 2019, 247).

The organization proposed by the collaborative would be an independent and international standard coordination body. Members would define priorities for technological standards, terminologies and guidelines to allow free flow of data under the individuals' control. The approach would combine a horizontal view with expert work groups (technical, design, legal, business, and so on), a sectoral approach with sector hubs (mobility, finance, health, administration, retail, and so on) and a cross-sectoral group. A technical board would coordinate the hubs and work groups with other standards organizations, legislators, regulators, academics, users, and so on (ibid., 252).

Robert Fay, Global Economy director at CIGI, has also recently suggested the creation of a data governance body with a broad mandate. In a recent essay entitled "Digital Platforms Require a Global Governance Framework," he proposed a new organization structured like the Financial Stability Board (FSB). The FSB, created after the 2008 financial crisis, was given a mandate by the Group of Twenty to "promote the reform of international financial regulation and supervision" with a role in standard

setting and in promoting members' implementation of international standards (Fay 2019, 28).

Fay's proposed Digital Stability Board (DSB) would be composed of a plenary body, which would set objectives and oversee the work of the board. It would consist of officials from countries that initially join the organization. It would work with standard-setting bodies, governments and policy makers, regulators, civil society and the platforms themselves via a set of working groups with clear mandates that would report back to the plenary. Funding would come from its member countries alongside voluntary donations and in-kind contributions via participation in the DSB working groups. It could report to the International Grand Committee on Big Data, Privacy and Democracy (IGC). The IGC, made up of a diverse set of 12 countries and more than 400 million citizens, has been active in investigating the behaviour of the FAAMG platforms, including their role in disseminating fake news.

DSTF

Given the need for data standards coordination to enhance digital cooperation, the lack of a clear mandate to create one under the UN umbrella and the fragmentation of standardization activities related to big data analytics, this paper proposes the creation a new institution that could be named the DSTF. Reporting to the proposed DSB plenary, it would be similar in structure to the IETF. The DSTF would be entrusted with a dual mandate: enabling the development of technical standards to create data value chains, and being accountable for the development of data governance standards to properly frame data collaboration platforms and the FAAMG platforms. The ultimate objective of the DSTF would be to create the required architecture for a "single data zone" where data can circulate freely between participating jurisdictions through a series of data collaboration platforms.

Digital cooperation will involve the creation of complex data value chains. Just as with traditional supply chains for tangible products, each segment of a given data value chain 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.

It should be anticipated that many standards and specifications will be required to properly frame data value chains. Codes, standards, guidelines, best practices and model technical regulations will be required to cover both the technical and governance layers.

The structuring of the DSTF would need to reflect the new realities of the digital age. Classical forms of governance do not apply. Technology moves so fast that by the time decision makers gather to prepare, discuss, approve, ratify and implement a convention or new agreement, the landscape has changed entirely. Analogue policy making will not work in a digital world. In order to be responsive, the DSTF would need to develop standards in a shorter time frame than the two to three years generally required in traditional standard-setting bodies. Once developed, some of the standards could be expected to be "evergreen," that is, to be updated on an ongoing basis in order to reflect new technologies and approaches and remain relevant. Traditional standard-setting bodies require a published standard to be reviewed every five years.

Ontology, Semantics, Definitions and Terminology

When industrial sectors were mostly vertical in nature, SDOs developed standards in silos. As a result, a multiplicity of domain-specific semantics, including product terminology, classification and properties were created and maintained, sometimes for many decades. 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. Big data analytics are, by definition, higher-level functions and will need to be based on a common ontology. It is a prerequisite for interoperability.

The DSTF would create a working group to lead the development and adoption of the right set of foundational standards covering ontology, semantics, definitions and terminology. These would be used by other working groups to ensure consistency across data value chains. They could also be used by regulators within the single data zone.

Technical Standards for Data Value Chains

The internet and the World Wide Web will provide the infrastructure backbone on which data value chains will be built. As outlined in a

recent CIGI paper entitled *Standards for the Digital Economy,* data value chains are composed of three segments: data collection and grading; data access, exchange and storage; and data analytics and solutions (Girard 2019b). Detailed standards, specifications and guidance are needed to achieve interoperability and make it possible for data collected in one data collaboration platform to be used by another within the single data zone.

As a first task, the DSTF would constitute working groups to articulate the roles and responsibilities associated with each of the three segments of a typical data value chain. It would identify standardization needs, adopt or adapt appropriate standards that have been developed, identify gaps and coordinate the development of new standards to fill these gaps. The task force could opt to mandate a limited number of existing bodies, such as the IETF, to develop the required technical standards and specifications. It could also issue requests for proposals and select appropriate SDOs to develop standards on its behalf. Once developed, technical standards and specifications would be adopted by the DSTF and added to a registry of compliant standards. As technology evolves quickly, data value chain standards should be updated as needed and a preference may be given to organizations that commit to evergreen standards.

Data Collection and Grading

Digital cooperation projects will require data from a multiplicity of sources to be successful. Existing data sets in analog format will be used in addition to digitized data sets stored in various databases in a multiplicity of formats. Traditional organizational data originates from various operational systems such as enterprise resource planning, customer relationship management, finance and human resources systems of record. In addition to traditional data sets, digital collaboratives will increasingly rely on streaming data from IoT devices, industrial sensors, cameras, clickstreams, servers and user app activity.

Metadata standards will be required to provide information about the characteristics of the data collection apparatus and about data set attributes, in order to precisely describe the features of available data sets; categorize and apply a grade to the data to make inferences about its quality; and label data sets and ensure they are tagged with appropriate IP and copyright mechanisms for traceability.

Data Access, Sharing, Exchange and Storage

This second segment of the data value chain is needed to make data accessible. It will serve as the interface to connect data sets with data users. New data collaboration platforms will be created to manage and track data flows on behalf of the participants making data available. They will also manage data access for AI and machine-learning organizations looking to generate new insights.

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. Standards will be required to describe and frame these different data access methods.

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 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). Standards will be needed to set performance requirements of APIs to be used in the single data zone and ensure interoperability between platforms.

Operators of data collaboration platforms will need to manage four core functions: data integration; systems interoperability; data provisioning; and data quality control. Tasks will 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, managing connections with other APIs, and reporting 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 Analytics and Solutions

This third segment of activities will be undertaken by a number of organizations from civil society, governments, academic and research organizations, and small and medium-sized enterprises engaged in AI and machine learning. Analytics functions could operate in a central location in an "IoT lab" in order to foster collaboration between participants. They could also operate in a decentralized way where each organization negotiates appropriate access rights to data 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.

Organizations engaged in data analytics will need standards to ensure that algorithms and solutions respect applicable regulations and ethical guidelines and are seen as trustworthy.

Data Governance Standards

Advances in digitization allow organizations to gather and store more data, enabling smarter and quicker decisions, but they are also giving rise to a new series of issues. How do organizations collect and distribute the right data at the right time? How should organizations deal with data ownership and copyright? How should personal information be treated? What rules should organizations follow regarding data residency? What are acceptable practices for the use of automated decision systems relying on AI?

Although some of these issues can be handled solely by organizations, many are framed by governments through enabling laws, regulations and policies. As regulators are not equipped to keep pace with rapid technological advancement, the DSTF would create a series of standards committees to develop and maintain the necessary foundational data governance standards. These standards would frame how digital cooperation initiatives and big tech platforms operate in the single data area. The standards committees would be composed of representatives from governments, industry, civil society and academics. They would operate through the established standards development process.

Jurisdictions participating in the DSTF should follow an approach similar to the one in place in the European Commission whereby they are accountable to make *standardization requests* to the DSTF regarding data governance issues that should be standardized, giving a clear mandate to the DSTF to launch standards committees as needed. Regulators would be called on to play a central role in the development of data governance standards to ensure they meet their policy objectives. Once a standard is developed, there would then be a *presumption of conformity* on the part of governments to adopt the standards and incorporate them by reference in regulations. This process would establish minimal standards that nation-states would agree to adhere to, which is paramount in order to establish a single data area.

Data governance issues that will likely require standardization support include:

- guidance on asserting ownership/IP/copyright over data collected by organizations and on tagging and tracking data use in the single data area;
- guidance on data valuation for the purposes of financial reporting and taxation;
- ensuring compliance to relevant privacy/digital identity requirements;
- ensuring compliance with relevant labour rights;
- facilitating personal data portability between data collaboration platforms;
- ensuring compliance with human rights regulations and requirements;
- establishing and enforcing relevant safe use, ethics and trustworthiness principles to data analytics, AI, machine learning and solutions;
- choosing and applying appropriate cyber security controls to data collection/access/analytics architecture, which may include encryption;
- defining appropriate data sovereignty and residency requirements to meet requirements under the single data area; and
- guidance on appropriate professional credentials and accountability for chief data officers accountable for data governance in organizations as well as data engineers (responsible for data collection and grading), data controllers (responsible for data access, sharing, exchange and storage), data scientists (responsible for AI, machine learning and algorithms) and data valuation professionals and data auditors.

Online Standards Registry

As indicated earlier, standardization activities in the ICT sector are highly fragmented, to the point where there is no pathway to quickly access relevant standards and specifications related to a specific domain. One key function of the DSTF would be to build and maintain a comprehensive registry of standards and normative documents covering data value chains and data governance. This corpus is needed for working groups to adopt/adapt existing standards, identify gaps and determine which standards bodies are best situated to undertake work on behalf of the DSTF.

A possible path forward to facilitate the upkeep of the proposed registry would be to use e-evidence processes to generate listings of relevant documents that could be reviewed by working groups. By inviting ICT SDOs, consortia and open-source software development platforms to share data sets, a knowledge base could be developed from machine reading of unstructured text and presenting relevant concepts through linked graphs of objects or relationships. Data sets would include both published codes, standards and specifications as well as new work items currently under development. Companies engaged in natural language processing such as Nuix and Analytics360 have developed methods to support research in sectors such as patent filing, litigation or university research publications.[37]

Machine-readable Standards

One of the major challenges of standards setting to support big data analytics will be to develop normative documents that can be read and interpreted by machines. Machine-readable standards will be necessary to introduce efficiency and automation in the standards development process and the development of new approaches for the certification of intangibles by digital certification processes involving algorithms (CSA Group 2019, 42). In addition, a growing number of nation-states have begun to transfer legislation and regulation texts into machine-readable formats. Providing machine-readable standards would be in line with that trend.[38]

Looking Forward

In the absence of a UN-mandated organization to coordinate the development of data value chains and data governance standards, nation-states sharing similar values can create a mechanism to enhance digital cooperation. The proposed DSTF could set the stage for the creation of a single data area among participating jurisdictions. Standardized data collaboration platforms, sourced with multiple data sets from participating jurisdictions, could shed new insights on persistent problems and benefit humanity. With the right data standards and compliance mechanisms in place, consumers can regain trust in global tech platforms, starting with FAAMG, and support the deployment of ethical AI. Finally, credible data standards can be incorporated by reference in national regulations. This will help avoid an unwieldy patchwork of rules and regulations.

Although it may sound counterintuitive, when it comes to setting global standards, leaders of the FAAMG platforms are often standing in the way of progress. They need to rethink their strategy regarding global voluntary standards setting. Stakeholders and shareholders do not want (and will not stand for) a patchwork of unenforceable, company-specific data governance policies, such as Google's proposed industry "standard" on data collection and digital advertising. Through the proposed DSTF, FAAMG platform leaders could join others in developing one suite of data governance standards. All players need to work toward "one standard, one test" in order to reap the benefits of big data analytics while managing unintended consequences.

37 See www.nuix.com/solutions/ediscovery; www.colorado.edu/oda/analytics360.

38 See https://digital-legislation.net/.

Works Cited

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.

Bremmer, Ian. 2019. "The End of the American Order: Ian Bremmer speech at the 2019 GZERO summit." November 18. Eurasia Group. www.eurasiagroup.net/live-post/end-of-american-order-ian-bremmer-2019-gzero-summit-speech.

Cihon, Peter. 2019. *Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development*. April. Future of Humanity Institute, University of Oxford.

CSA Group. 2019. "The Digital Age: Exploring the Role of Standards for Data Governance, Artificial Intelligence and Emerging Platforms." May. Mississauga, ON: CSA Group.

CSTD. 2015. "Mapping of international Internet public policy issues." April. Geneva, Switzerland: UNCTAD. https://unctad.org/meetings/en/SessionalDocuments/ecn162015crp2_en.pdf.

Darabi, Anoush 2018. "New Zealand explores machine-readable laws to transform government." Apolitical, May 11. https://apolitical.co/solution_article/new-zealand-explores-machine-readable-laws-to-transform-government/.

Digital Cooperation. 2019. *The Age of Digital Interdependence. Report of the UN Secretary-General's High-level Panel on Digital Cooperation*. June 10. www.un.org/en/pdfs/DigitalCooperation-report-for%20web.pdf.

Fay, Robert. 20119. "Digital Platforms Require a Global Governance Framework." In *Models for Platform Governance*. Waterloo, ON: CIGI. www.cigionline.org/publications/models-platform-governance.

Girard, Michel. 2019a. *Big Data Analytics Need Standards to Thrive: What Standards Are and Why They Matter*. CIGI Paper No. 209. www.cigionline.org/publications/big-data-analytics-need-standards-thrive-what-standards-are-and-why-they-matter.

———. 2019b. *Standards for the Digital Economy: Creating an Architecture for Data Collection, Access and Analytics*. CIGI Policy Brief No. 155. www.cigionline.org/publications/standards-digital-economy-creating-architecture-data-collection-access-and-analytics.

Guterres, António. 2019. "Time for global cooperation on governing digital technology." *Nikkei Asian Review*, June 26. https://asia.nikkei.com/Opinion/Time-for-global-cooperation-on-governing-digital-technology.

Iny, Yoni. 2019. "The Modern Data Lake Architecture: Unlocking the value of streaming data by simplifying big data discovery, processing and management in cloud data lakes." Technical Whitepaper. Upsolver.

Leblond, Patrick and Susan Ariel Aaronson. 2019. *A Plurilateral "Single Data Area" Is the Solution to Canada's Data Trilemma*. CIGI Paper No. 226. Waterloo, ON: CIGI. www.cigionline.org/publications/plurilateral-single-data-area-solution-canadas-data-trilemma.

Marr, Bernard. 2018. "How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read." *Forbes*, May 21. www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#73555d5660ba.

McClelland, Calum. 2018. "What Is an IoT Platform?" IoT for All, August 17. www.iotforall.com/what-is-an-iot-platform/?utm_source=CampaignMonitor&utm_medium=email&utm_campaign=IFA_Signup_Welcome_Articles.

Plautz, Jason. 2018. "Cheap, Portable Sensors are Democratizing Air-Quality Data." *Wired*, November 7. www.wired.com/story/cheap-portable-sensors-are-democratizing-air-quality-data/.

Privacy Tech. 2019. "Une nouvelle gouvernance pour les données au XXIe siècle : des standards pour la circulation et la protection des données personnelles." April. Paris, France: Privacy Tech.

UNCTAD. 2019. *Digital Economy Report 2019*. New York, NY: United Nations. https://unctad.org/en/PublicationsLibrary/der2019_en.pdf.

UN General Assembly. 2002. "World Summit on the Information Society." January 31. A/RES/56/183.

WHO. 2019. *Global Strategy on Digital Health 2020–2024*. March 26. https://extranet.who.int/dataform/upload/surveys/183439/files/Draft%20Global%20Strategy%20on%20Digital%20Health.pdf.

WMO. 2017. *WIGOS Metadata Standard*. Geneva, Switzerland: WMO. https://oscar.wmo.int/surface/#/.

Centre for International
Governance Innovation

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

@cigionline