



Discerning signal from noise:

The state of global AI standardization and what it means for Canada



TABLE OF CONTENTS

FRONT MATTER	1
Executive summary	2
INTRODUCTION	4
PART I: POLICY CONTEXT AND KEY INITIATIVES	6
Section I: States and regions	13
Section II: International standards organizations	16
PART II: ANALYSIS	18
Section I: Introduction	18
Section II: Key international actors	18
Section III: The role of standards and conformity assessment	19
Section IV: Consequences of a fragmented international landscape	20
PART III: RECOMMENDATIONS	23
Section I: Introduction	23
Section II: General recommendations	28
Section III: Opportunities for bilateral activities	30
Section IV: Priority use cases for AI standardization	31
CONCLUSION	33
APPENDIX I: METHODOLOGY	34
APPENDIX II: CONSULTATIONS	35

Prepared for the Standards Council of Canada by the <u>Schwartz Reisman Institute for Technology and Society</u> at the University of Toronto, with contributions from INQ Law, January 2023.







FRONT MATTER

Executive summary

This white paper aims to equip the Standards Council of Canada's (SCC) Al and Data Governance Standardization Collaborative (Collaborative) with the information necessary to understand how best to leverage artificial intelligence (AI) standards and conformity assessment (as well as various assurance techniques mentioned hereafter) to advance Canada's interests in this space, both at home and abroad. The points made in this paper come from interviews conducted by the policy team at the Schwartz Reisman Institute for Technology and Society, and they are supported by secondary research.

To do so, it first considers standards initiatives of general application; focuses on prominent national and international policy initiatives of greatest relevance to the Canadian market; and discusses the risks and opportunities associated with such initiatives, such as international standards fragmentation and bilateral agreements. In examining these risks, it analyzes the key actors in international standardization efforts: the United States, China, and the European Union, all of whom have large domestic markets and a considerable ability to influence the international Al landscape. It notes that, despite a diverging regulatory landscape, assurance techniques have emerged as a strong tool for international cohesion, as each of the regions examined has signalled intent to utilize assurance techniques as part of its Al strategy. These insights were collected through both primary research (interviews) and secondary research (including journal articles and legal documents). See Appendix I for additional details.

Equipped with this information, the Collaborative is invited to consider a series of recommendations related to Canada's strategy on AI standardization. The recommendations focus first on opportunities for bilateral cooperation between SCC and the major standards bodies in the U.S. and EU. General recommendations are then laid out, including recommendations on increasing Canadian participation in developing international AI standards, and enhancing Indigenous and small- and medium-sized enterprise participation in these processes. The importance of meaningful engagement with Canadian Indigenous governments is stressed, and the issue of Indigenous data sovereignty raised as one that must be grappled with. The Collaborative is urged to focus on enhancing openness, accessibility, and diversity in standards development work.

Priority standards and conformity assessment programs are identified for the Collaborative's increased attention, and the Collaborative is invited to look globally for inspiration on AI standardization initiatives, with several initiatives highlighted. Additional recommendations look at enhancing the relationship between industry, with greater resources, and academics, with particular research and insights, and creating new ways for these actors to bring the results of these collaborations to technical standardization committees without needing to sit in committee seats.

Finally, the Collaborative is presented with priority use cases for AI standardization. Selection criteria for identifying priority use cases in this field are identified and explained, and example use cases put forward.

Key terms¹

Standards are documents that set out established practices arrived at by consensus and approved by a recognized body. They provide for common and repeated use, rules, guidelines, or characteristics for activities or their results, and are aimed at achievement of the optimum degree of order in a given context.² Standards are typically voluntary but can become mandatory when enforced by laws or regulations, for example, for health or safety reasons.³ Standardization is the development and application of standards. It includes:

- the work of the committees that develop standards;
- the publication of standards by standards development organizations;
- the recognition of standards by national standards bodies such as SCC;
- the application of standards by businesses, suppliers and customers;
- the verification that products or services conform to applicable standards (conformity assessment);
- the accreditation of organizations that provide conformity assessment services;
- the use of standards and conformity assessment as an element in public policy as well as in international trade.

Assurance ecosystems, or ecosystems of trust, are made up of several different components and are intended to provide consumers with justified trust in a particular product or service. For example, the UK has laid out a roadmap to an effective AI assurance ecosystem, detailing the need for third-party auditors, certification, assessments, and regulation to create a balanced ecosystem in which consumers can trust that any AI systems in use have met a certain safety threshold.⁴ Although proposals for assurance ecosystems may differ, standards and conformity assessment are a vital component of any such ecosystem.

Impact assessments evaluate the impact a particular activity or system could have. For example, an impact assessment of an AI system that decides who will receive a loan might identify whether and how much those seeking loans might be affected. Impact assessments for AI may build off existing impact-assessment frameworks in fields such as environmental protection, human rights, or data protection.⁵

Conformity assessment confirms whether a service, system, or product adheres to the requirements of a particular standard. Certification bodies are often accredited by Accreditation bodies in their country. Accreditation bodies typically are members of the International Accreditation Forum (IAF) and/or the International Laboratory Accreditation Cooperation (ILAC). Accreditation provides independent confirmation of competence of a certification body. For example, SCC is Canada's leading accreditation organization, ensuring conformity assessment bodies meet the highest national and international standards against nationally and internationally recognized standards.

Certification is a process through which an independent body attests that an organization or its personnel, systems, or products meet objective standards of quality or performance, typically through the issuance of a "mark" or "label."

¹ Although the following definitions are generally aligned with those of ISO/IEC, please note that minor differences may remain.

² "Types of Standards," Standards Council of Canada, online (scc.ca).

³ "Types of Standards," scc.ca.

⁴ "The roadmap to an effective AI assurance ecosystem," Centre for Data Ethics and Innovation, 8 Dec. 2021, online (gov.uk).

⁵ "Recommendation: develop baseline model for Al impact assessments," *Project Sherpa*, online (<u>project-sherpa.eu</u>). Note that there is a lack of clarity in the nomenclature about impact and risk assessments. While many use the terms interchangeably, some distinguish the two concepts. Those in the latter camp denote risk assessments as an exercise undertaken pre-deployment to understand potential risks, and impact assessments as a task performed *ex post* to assess actual impact.

Risk-based legislative approaches emphasize how much risk might arise for a particular product or system and regulate accordingly, with greater obligations placed on providers of high-risk products or systems. The U.S., Canada, and the EU have signaled such an approach to regulating AI, although Canada uses the language of "impact" rather than "risk." ⁶

Prescriptive legislative approaches refer to initiatives that are specific and highly detailed. They prescribe certain rules to follow or activities to carry out. The EU has signaled intent to take a prescriptive approach to regulating AI.

Principles-based legislative approaches are less definitive, enumerating principles to follow without specifying precise activities or requirements to undertake. In the context of AI, such principles are often drawn from the Organization for Economic Co-operation and Development (OECD) AI Principles. Both the UK and the U.S. have introduced initiatives that signal a principles-based approach to governing AI, although neither seem poised to encode this approach in formal legislation. B

The **OECD AI Principles** are a set of guiding principles to guide actions and policy decisions related to AI. Since their adoption, these principles have heavily influenced major national documents and initiatives related to AI. 10

National Mirror Committees are national structures set up to mirror the work of international standardization organizations (e.g., ISO/IEC) for Standardization Technical Committee within a national standards body. A National Mirror Committee brings together stakeholders to establish a national consensus position on a given standard. The Standards Council of Canada is the national standards body for Canada.

Civil-society groups are non-governmental groups with an organized structure, particular goals, and a mission statement that typically centres around protecting the interests of group members, including efforts to minimize or eradicate systematic harms.

⁶ "Bill C-27, Part 3 Artificial Intelligence and Data Act," *House of Commons of Canada*, 16 June 2022, online (parl.ca); "Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (AI Act)," *European Commission*, 21 Apr. 2021, online (eur-lex.europa.eu). At the time of writing, the most recent draft can be found here: artificialintelligenceact.eu. Newer drafts are posted here as they become available: artificialintelligenceact.eu.

⁷ "OECD AI Principles overview," Organisation for Economic Co-operation and Development (OECD), online (oecd.ai).

⁸ Prescriptive approaches should not be assumed to be more effective than principles-based approaches; increased specificity does not necessarily translate into greater effectiveness at regulating AI.

⁹ "OECD AI Principles overview," <u>oecd.ai</u>.

¹⁰ For example, they are reflected in a recent UK white paper, the U.S.'s risk-management and rights frameworks for AI, and the EU's AI legislation.

INTRODUCTION

Artificial intelligence (AI) is rapidly changing the landscape in every sphere of society today. While governments across the world have been rushing to catch up, there remains a lag in the development and adoption of appropriate policy instruments that can address AI's unique challenges. However, some progress has been made, and Canada can position itself to seize on these opportunities.

In 2017, Canada became the first country in the world to release a national AI strategy. ¹¹ Since then, it has continued to earn a position on the world stage when it comes to AI, ranking first in the world for government strategy and commitment to AI projects, fourth for its overall AI ecosystem, and tenth for research and development. ¹² Canada is fully engaged in international standardization efforts, and is establishing an AI and Data Governance Standardization Collaborative (Collaborative). ¹³ The Collaborative is a multi-stakeholder forum coordinated by the Standards Council of Canada (SCC) that will identify priorities for the development and adoption of AI and Data Governance standards and conformity assessment approaches, including strategic programs and initiatives. This will catalyze the realization of Canada's objectives in this rapidly evolving space.

It has been estimated that AI will add 15 trillion USD to the global economy by 2030, and Canada is well situated to share in these benefits. ¹⁴ In the next decade, the AI industry is predicted to add 16 billion CAD to the Canadian economy and create more than 16,000 jobs. ¹⁵ In fact, from 2015 to 2020, Canada's AI sector attracted 3 billion CAD in investment and created 50,000 jobs. ¹⁶ These advances are growing rapidly. In 2019, Canadian AI companies received 57 venture-capital deals worth 658 million USD, an increase of almost 50% from 2018. ¹⁷ With the highest concentration of AI start-ups in the world, and the establishment of AI-specific research and development centres in large cities by multinational enterprises, the economic opportunities for Canada are clear. ¹⁸ Realizing these opportunities will require a defined and well-informed strategy on AI standards and conformity assessment.

Standards and conformity assessment play an important role in creating cohesive, common frameworks upon which new technologies can be built, tested, and safely deployed. They offer increased consumer trust and safety, protecting not only the interests of consumers and civil-society groups, but also the interests of businesses that need trust and safety in order to sell their products or services and protect and promote various interests. They facilitate innovation, interoperability of products and services across international borders, market access, and competition. These benefits remain salient for AI; they will be vital for ensuring the same levels of trust, safety, innovation, and interoperability that we have come to expect from other regulated products and services.

Beyond general international trade considerations, standards and conformity assessment also present a clear opportunity to affirm and export values. Accreditation, and being a signatory to a Multilateral Recognition Arrangement through IAF or ILAC, significantly impacts trade for companies that use conformity assessment to reach new markets¹⁹. Technical requirements are a top concern on trade barriers. The objective of the World Trade Organization Technical Barriers to Trade (WTO TBT)

^{11 &}quot;Pan-Canadian Artificial Intelligence Strategy," Innovation, Science, and Economic Development Canada, online (ised-isde.canada.ca).

^{12 &}quot;The Global AI Index," Tortoise Media, online (tortoisemedia.com).

^{13 &}quot;Standardization to play a key role in the modernization of Canada's regulatory system," Standards Council of Canada, 22 June 2021, online (scc.ca).

¹⁴ Holmes, Frank. "AI Will Add \$15 Trillion To The World Economy By 2030," Forbes, Great Speculations, 25 Feb. 2019, online (forbes.com).

¹⁵ Quiroz, Karicia and Anam Elahi. "Canada's Advantage for Al: An Ecosystem Fueled by Talent and Innovation," *Invest in Canada*, Sept. 2020, online (investcanada.ca).

¹⁶ "Canada's AI Ecosystem: Government Investment Propels Private Sector Growth," University of Toronto, online (gro.utoronto.ca).

¹⁷ "Canada's Advantage for AI," <u>investcanada.ca</u>.

^{18 &}quot;Canada's AI Ecosystem," gro.utoronto.ca.

¹⁹ "The effects of cooperation in accreditation on international trade," ScienceDirect, online (https://www.sciencedirect.com/science/article/pii/S092552731830063X)

Agreement is to support the reduction of barriers to trade by enhancing cooperation between accreditation systems, preventing the preparation, adoption and application of different technical regulations and conformity assessment procedures that are unreasonable. ²⁰ Thus, the international AI standards arena presents significant challenges, not least due to the necessity of achieving a certain level of international alignment amidst a complex and fragmented series of initiatives and actors. As ever, states and regions with the largest markets have considerable influence and resources when it comes to creating standards and accessing "assurance" ecosystems that meet their needs and the needs of key trading partners. States with smaller markets must look at the requirements for access to global markets, including the countries that shape them, when considering their standardization initiatives.

Over the past few years, national and international standards development organizations have undertaken a variety of initiatives to promote research, foster innovation, and accelerate the commercialization of AI products and services. More recently, national governments have added AI to their legislative agendas. Despite divergent approaches, each has committed to exploring AI regulatory models that rely, in different ways, on adherence to industry standards through conformity assessment such as through accreditation bodies that oversee audits, certification schemes, and other assurance tools.

However, a comprehensive set of international AI standards and conformity assessment has yet to fully emerge. In this constantly evolving landscape, governments, businesses, researchers, and the public are faced with the considerable challenge of scanning and understanding an increasingly complex international AI standards landscape, while attempting to seize the opportunities it presents.

For government stakeholders, a fulsome understanding of the state of play presents major opportunities and the potential for competitive advantage. In particular, it allows them to influence the creation of international AI standards, use these standards as the agile tools necessary to regulate fast-moving AI technologies, ensure interoperability and market access to protect domestic growth in this sector, and protect and promote shared norms, assurance ecosystems and values.

For industry, it will be valuable to know what is coming for AI standardization, to understand how to contribute to standardization initiatives, to leverage standards for regulatory compliance, and to remain adaptable and capable of accessing national and international markets through accreditation systems.

For academics, greater awareness of the growing importance of AI standards and conformity assessment can provide opportunities to contribute to the development of technical standards through their work in domains such as AI ethics, safety, alignment, socio-technical approaches to measuring bias, and techniques for evaluating large and powerful systems as they emerge.

For the general public and, in particular, for communities at greater risk of experiencing harm from AI deployment, understanding how AI standards and conformity assessment are being developed and playing a meaningful role in these processes is critical.²¹ Doing so is vital to the quality, credibility, and democratic legitimacy of these instruments as they take on greater importance in both the regulated and unregulated environments that impact our daily lives. This is particularly true for Indigenous peoples in Canada, for whom deployment of AI poses several risks.

²⁰ Technical Information on Technical Barriers to trade, Work Trade Organization, online (https://www.wto.org/english/tratop_e/tbt_e/tbt_info_e.htm)

²¹ However, it will be vital for the Collaborative to approach involving the general public with caution, given clear research to show that minority groups are rarely involved in these processes due to systemic barriers. The need to address this is discussed in further detail in the recommendations section.

PART I: POLICY CONTEXT AND KEY INITIATIVES

Section I: States and regions

Canada

Canada was the first country in the world to develop a national AI strategy when the Canadian Institute for Advanced Research released the Pan-Canadian Artificial Intelligence Strategy (PCAIS) in 2017. PCAIS lays out Canada's goals of being a world leader in AI and is split into three major pillars: commercialization, standards, and talent and research. Phase 1 of PCAIS focused on building up Canadian talent and an AI ecosystem through activities such as establishing research centers and unifying Canada's three major national AI institutes: Alberta Machine Intelligence Institute, Mila – Quebec Artificial Intelligence Institute, and the Vector Institute for Artificial Intelligence. Phase 2 of PCAIS now seeks to form connections between AI talent, cutting-edge research, and commercialization and adoption. 23

Under Phase 2 of Federal Budget 2021, the SCC received 8.6 million CAD to advance the development of AI standards and conformity assessment. Early initiatives under this funding include the launch of a unique pilot to define and test requirements for a conformity assessment program for AI management systems (AIMS pilot). The first stage of the pilot includes one conformity assessment body (under the accreditation of SCC) and one AI developer or user assessing against the International Organization for Standardization draft standard requirements for AI management systems (ISO/IEC DIS 42001, hereinafter "AIMS"), as well as the Canadian government's Algorithmic Impact Assessment.²⁴ The pilots are being launched in collaboration with the Responsible AI Institute (RAII), which provides a research-based framework for key use cases.²⁵

The SCC Innovation Initiative was launched in 2018 with the aim of financially supporting Canadian innovators who hoped to develop tailored standardization strategies.²⁶ The Initiative planned to support the development of 63 strategies by the end of 2022. Previous efforts by the Data Governance Standardization Collaborative (DGSC) have highlighted the potential for the SCC Innovation Initiative to increase access to and participation in standardization activities by small- and medium-sized enterprises (SMEs).²⁷ The DGSC's Data Governance Standardization Roadmap also includes recommendations relevant to the Collaborative as it begins work on AI standardization.²⁸ The section of this document addressing Indigenous data sovereignty is noteworthy, as will be discussed further in this paper.

Innovation, Science and Economic Development Canada launched the Superclusters Innovation Initiative in 2017 (recently renamed the Global Innovation Clusters), which intends to spur growth and innovation in some of Canada's most promising industries by encouraging collaboration between industry and academia and providing up to approximately 1 billion CAD in funding. One of Canada's five Clusters is Scale AI, focused on AI and supply chain technology, with the goal of making Canada a global leader in exports and ensuring Canadian products are first to market.²⁹

²² "Pan-Canadian Artificial Intelligence Strategy," <u>ised-isde.canada.ca</u>.

²³ "Government of Canada launches second phase of the Pan-Canadian Artificial Intelligence Strategy," *Innovation, Science, and Economic Development Canada*, 22 June 2022, online (canada.ca).

²⁴ "SCC launches accreditation pilot for AI Management Systems," Standards Council of Canada, 10 June 2022, online (scc.ca).

²⁵ "SCC launches accreditation pilot for Al Management Systems," <u>scc.ca</u>.

²⁶ "The Innovation Initiative," Standards Council of Canada, online (scc.ca).

²⁷ "The Innovation Initiative," scc.ca.

²⁸ "Canadian Data Governance Standardization Roadmap," Standards Council of Canada, 28 June 2021, online (scc.ca).

²⁹ "Global Innovation Clusters," Innovation, Science, and Economic Development Canada, online (ised-isde.canada.ca).

Tabled as part of *Bill C-27 Digital Charter Implementation Act, 2022,* Canada's recently proposed *Artificial Intelligence and Data Act* represents the federal government's first comprehensive attempt at regulating AI.³⁰ This proposed legislation takes a risk-based approach and, if passed, persons responsible for the development or deployment of AI that has been assessed as "high-impact" will have to comply with more stringent obligations, many of which will be articulated in regulations issued under the law should it ultimately be adopted. While the preamble to Bill C-27 states that, "the design, development and deployment of artificial intelligence systems across provincial and international borders should be consistent with national and international standards to protect individuals from potential harm," it remains unclear whether and how Canada—like some of its allies and trading partners described below—will rely on AI standards and conformity assessment as part of an emerging regulatory regime.

European Union

The EU's proposed *Artificial Intelligence Act* (*AI Act*) has had a transformative impact on international AI regulation and policy and the role of standards therein. The draft legislation takes a risk-based, prescriptive approach, grouping AI into three categories: unacceptable risk, high-risk, and low or minimal risk. AI in the first category will be completely prohibited, whereas AI in the high-risk category will be subject to certain requirements. Low-risk AI will be exempt from regulation, ³¹ and general-purpose AI will be subject to a smaller set of tailored obligations. ³²

The general obligations applicable to both high-risk and general-purpose systems include: the implementation of risk-management systems and practices to ensure transparency, security, maintenance of technical documentation, and data quality; accurate, robust, and secure systems; and post-market monitoring.³³ All organizations developing, offering, or making use of high-risk AI will need to undergo conformity assessment and implement practices that enable traceability, human oversight, and quality management.³⁴ These obligations will be mandatory and specifically set out in the legislation (as opposed to general principles open to interpretation, such as safety and security).³⁵

Other key pieces of legislation include the *Digital Markets Act* (2022) and the *Digital Services Act* (2022). ³⁶ Their joint goals are to create a safer digital space in which the fundamental rights of all users of digital services are protected and to establish a level playing field to foster innovation, growth, and competitiveness within and outside of the EU. ³⁷ While these goals indirectly impact the data-sharing and competition functions of AI, the *Digital Services Act* also regulates AI directly by imposing transparency obligations on recommender systems that drive optimized access to information. ³⁸

With respect to standards, the AI Act will rely on compliance with harmonized standards as a means to demonstrate conformity with its requirements. The European Commission will instruct the European Committee for Standardization (CEN) and the European Committee for Electrotechnical Standardization (CENELEC) to prepare these harmonized standards, with

³⁰ "Artificial Intelligence and Data Act," parl.ca.

³¹ This will only be the case if the AI provider has explicitly excluded all high-risk uses in the instructions of use or information accompanying the general purpose AI system (see AI Act, Draft for COREPER, Article 4c, artificialintelligenceact.eu).

³² As of the most recent Presidency compromise draft. However, it is important to note that general purpose AI may be considered high-risk on its own or in collaboration with other high-risk AI depending on the context of its functionality (see *AI Act*, Draft for COREPER, Preamble s. 12(c), artificialintelligenceact.eu).

³³ "Al Act," <u>eur-lex.europa.eu</u>.

^{34 &}quot;Al Act," eur-lex.europa.eu.

^{35 &}quot;AI Act," eur-lex.europa.eu.

³⁶ "Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services," *European Commission*, 15 Dec. 2020, online (eur-lex.europa.eu); "Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act) (Text with EEA relevance)," *Official Journal of the European Union*, 14 Sept. 2022, online (eur-lex.europa.eu).

³⁷ "The Digital Services Act Package," European Commission, 2022, online (ec.europa.eu).

³⁸ "Digital Services Act," Article 29, European Commission, 15 Dec. 2020, online (eur-lex.europa.eu).

the European Telecommunications Standards Institute (ETSI) likely playing a minor consulting role. ³⁹ In preparation, CEN and CENELEC established a joint technical committee ⁴⁰ in March 2021 to develop and adopt AI standards that address European needs and support EU policies and values, including international standards already in place or under development by bodies like ISO/IEC JTC1 SC42. ⁴¹ One of its initial projects is developing an *ad-hoc* committee to provide guidance on a conformity assessment that gauges compliance with the requirements in the *AI Act* for high-risk systems. ⁴² For its part, ETSI set up the Operational Co-ordination Group on Artificial Intelligence to coordinate AI standardization efforts handled by ETSI's technical bodies by identifying synergies, best practices, common activities, and gaps; proposing solutions; and evaluating the impact of AI on standards production. ⁴³

The result of this work will be a set of standards that, if followed, will help ensure compliance with the *AI Act*. A table from the European Commission's Joint Research Centre (Table 1) illustrates how certain international and European standards relate to key requirements for high-risk AI systems under the *AI Act*. ⁴⁴ As with all harmonized standards, the EU has indicated that adherence will not be mandatory. That is, regulated entities may take alternative routes to illustrate they have met the *AI Act*'s essential requirements, for example, by independently complying with its requirements using conformity assessment. ⁴⁵ A figure from CEN and CENELEC (Figure 1) outlines the planned coordination between European legislation and standard-making initiatives for AI. ⁴⁶



Figure 1: Planned coordination between European legislation and AI standard-making initiatives. Source: CEN and CENELEC.

⁴¹ Nonnecke, Brandie and Philip Dawson. "Human Rights Implications of Algorithmic Impact Assessments: Priority Considerations to Guide Effective Development and Use," *Harvard Kennedy School, Carr Center Discussion Paper Series,* 21 Oct. 2021, online (<u>carrcenter.hks.harvard.edu</u>); "Artificial Intelligence," *European Committee for Standardization* and *European Committee for Electrotechnical Standardization*, online (<u>cencenelec.eu</u>); "Work Programme 2022," *European Committee for Standardization* and *European Committee for Electrotechnical Standardization*, Jan. 2022, online (<u>cencenelec.eu</u>).

³⁹ "How the EU's Flawed Artificial Intelligence Regulation Endangers the Social Safety Net: Questions and Answers," *Human Rights Watch*, 10 Nov. 2021, online (hrw.org).

⁴⁰ CEN-CLC/JTC 21.

⁴² Ezeani, Gabriella, et al. "A survey of artificial intelligence risk assessment methodologies: The global state of play and leading practices identified," *Trilateral Research*, 27 Jan. 2022, online (trilateralresearch.com).

⁴³ "Artificial intelligence," *European Telecommunications Standards Institute*, online (<u>etsi.org</u>); Frost, Lindsay. "Standardization request in support of safe & trustworthy artificial intelligence," *European Telecommunications Standards Institute*, 1 June 2022, online (<u>etsi.org</u>); "Terms of Reference for OCG subgroup on Artificial Intelligence (OCG AI)," *European Telecommunications Standards Institute*, 21 Oct. 2021, online (<u>portal.etsi.org</u>).

⁴⁴ Nativi, Stefano and Sarah De Nigris, "Al Watch: Al Standardisation Landscape state of play and link to the EC proposal for an Al regulatory framework," Joint Research Centre, European Commission, 14 July 2021, online (standict.eu), pp. 20–21.

⁴⁵ "Al Act," <u>eur-lex.europa.eu</u>.

⁴⁶ Kohler, Constant and Laurens Hernalsteen. "Drafting Harmonized Standards in support of the Artificial Intelligence Act (AIA)," European Committee for Standardization and European Committee for Electrotechnical Standardization, 21 Mar. 2022, online (cencenelec.eu), p. 5.

Requirements	Data and data governance	Risk management system	Technical data and record keeping	Transparency and information to users	Human oversight	Accuracy, robustness, and cybersecurity	Quality management system
SDO							
ISO and ISO/IEC JTC1	ISO/IEC 25024; ISO/IEC 5259; ISO/IEC 24668	ISO/IEC 4213; ISO/IEC 25059; ISO/IEC 24029-2	ISO/IEC 5338; ISO/IEC 5469; ISO/IEC 24368; ISO/IEC 24372; ISO/IEC 24668	ISO/IEC 24027; ISO/IEC 24028; ISO/IEC 5338; ISO/IEC 24368; ISO/IEC 24372; ISO/IEC 24668; ISO/IEC 4213		ISO/IEC 24027; ISO/IEC 24028; ISO/IEC 24029; ISO/IEC 5469	ISO/IEC 23894; ISO/IEC 38507; ISO/IEC 42001; ISO/IEC 25059
IEEE	ECPAIS Bias; IEEE P7002; IEEE P7003; IEEE P7004; IEEE P7005; IEEE P7006; IEEE P7009; IEEE P2801; IEEE P2807; IEEE P2863	IEEE P7009; IEEE P2807; IEEE P2846	ECPAIS Transparency; IEEE P7000, IEEE P7001, IEEE P7006, IEEE P2801; IEEE P2802; IEEE P2807; IEEE P2863; IEEE P33333.1.3	ECP AIS Bias; ECP AIS Transparency; ECP AIS Accountability; IEEE P7000; IEEE P7001; IEEE P7003; IEEE P7004; IEEE P7005; IEEE P7007; IEEE P7008; IEEE P7009; IEEE P7011; IEEE P7012; IEEE P7014; IEEE P2863; IEEE P3652.1	ECP AIS Accountability; ECP AIS Transparency; IEEE P7000, IEEE P7006, IEEE 7010; IEEE P7014; IEEE P2863	ECPAIS Transparency; IEEE P7007; IEEE P7009; IEEE P7011; IEEE P7012; IEEE P2802; IEEE P2807; IEEE P2846; IEEE P2863; IEEE P33333.1.3	IEEE 2801; IEEE P2863; IEEE P7000
ETSI	DES/eHEALTH- 008; GR CIM 007; GS CIM 009; ENI GS 001; GR NFV- IFA 041; DGR SAI 002; TR 103 674; TR 103 675; TS 103 327; TS 103 194; TS 103 195.2, SAREF Ontologies	GS ARF 003; GR CIM 007; ENI GS 005; GR NFV-IFA 041; DGS SAI 003; EG 203 341; TS 103 194; TS 103 195.2; TR 103 821	DES eHEALTH- 008; ENI GS 005; DGR SAI 002; SAREF Ontologies; GR CIM 007; GS CIM 009	DES/eHealth- 008; GS CIM 009; DGR SAI 002; SAREF Ontology	DES/eHEALTH- 008; DGR SAI 005	GS ARF 003; GR CIM 007; ENI GS 001; ENI GR 007; DGR SAI 001; DGR SAI 002; DGS SAI 003; GR SAI 004; GS ZSM 002; TR 103 674; TR 103 675; TS 103 327; GS 102 181; GS 102 182	
ITU-T	ITU-T Y.3170; ITU-T Y.MecTa-ML; ITU-T Y.3531; ITU-T Y.3172; ITU-T H.CUAV- AIF; ITU-T F.VS-AIMC; ITU-T Y.4470; Y.Supp.63 to ITU-T Y.4000 series	ITU-T Y.qos- mi-are; ITU-T Y.3172; ITU-T H.CUAV-AIF; ITU-T F.VS- AIMC; ITU-T Y.4470		ITU-T Y.4470		ITU-T Y.3170; ITU-T Y.qos-mi- arc; ITU-T Y.MecTa-MI; ITU- T Y.3531; ITU-T Y.3172; ITU-T H.CUAV-AIF; ITU- T F.VS-AIMC; ITU- T Y.4470	

Table 1: Overall representation of mapped standards (already published standards are in bold). Source: Joint Research Centre, European Commission.

Some commentators have predicted the vast majority of organizations will choose to follow the harmonized standards. However, these harmonized standards engender potential pitfalls. For example, human rights organizations have raised concerns that insufficient consideration of human rights in harmonized standards may result in the *AI Act*'s stated objectives on human rights not being met through adherence to these standards alone. ⁴⁷ Others have expressed concerns that only large corporations—primarily large technology companies—will have the technical resources and capacity to meet the harmonized standards, questioning whether and how SMEs will be able to self-assess against these standards.

In an effort to better support SMEs, the European Commission, in collaboration with the government of Spain, launched the first regulatory sandbox pilot focused on AI in financial services in June 2022. ⁴⁸ Their goals are to promote the innovation of products and technologies that currently do not have a place in existing regulation, pave the way for the seamless integration of future regulation, and enable the creation of a pan-European system of AI-centric sandboxes. ⁴⁹ More tangibly, the Spanish sandbox intends to facilitate closer relationships between regulators and AI developers in order to define best practices, test the feasibility of provisions, provide guidance on compliance, and guide the implementation of the AI Act in all member states. Through this regulatory sandbox, the Spanish government and the European Commission expect to develop easy-to-implement best practice guidelines that will help companies (particularly SMEs and start-ups) implement rules to support their compliance with the AI Act. ⁵⁰

United Kingdom

The UK released its National AI Strategy in 2021, and a follow-up white paper, "Establishing a pro-innovation approach to regulating AI," in 2022.⁵¹ The UK plans to create a voluntary and principles-based regime that relies mainly on voluntary compliance with broad, cross-sectoral ideals.⁵² Although the National AI Strategy signalled a plan to pursue a horizontal regulatory framework, the white paper released a series of general AI principles to be followed and suggests that any regulation will be done on a sector-by-sector basis by existing regulators.⁵³

With respect to standards, the Alan Turing Institute, supported by the British Standards Institution and the National Physical Laboratory, recently launched the Al Standards Hub. The Hub aims to help Al organizations in the UK understand, develop, and benefit from international Al standards. It focuses on four pillars of activity: establishing an observatory, promoting community and collaboration, knowledge and training, and research and analysis. In so doing, the Hub will offer a single, central, and comprehensive information source for the UK's Al community, providing practical tools, educational materials, and a publicly available searchable database of Al standards.⁵⁴ The Alan Turing Institute describes the Hub as a complement to the UK's pro-innovation regulatory approach, and has said it should both improve the governance and unlock the economic benefits of Al.⁵⁵

Additionally, the Centre for Data Ethics and Innovation (CDEI) published a "roadmap to an effective AI assurance ecosystem," pointing to the necessity of an assurance ecosystem consisting of both standards and impact assessments, and conformity assessment tools to achieve consumer trust in AI.⁵⁶ In particular, the CDEI's roadmap highlights the emerging importance of

⁴⁷ "How the EU's Flawed Artificial Intelligence Regulation Endangers the Social Safety Net," hrw.org; "Human Rights Implications of Algorithmic Impact Assessments," carroenter.hks.harvard.edu.

^{48 &}quot;First regulatory sandbox on Artificial Intelligence presented," European Commission, 27 June 2022, online (digital-strategy.ec.europa.eu).

⁴⁹ Spanish Regulatory Sandbox website (sandboxspain.com).

⁵⁰ "First regulatory sandbox on Artificial Intelligence presented," <u>digital-strategy.ec.europa.eu</u>.

⁵¹ "Establishing a pro-innovation approach to regulating AI," UK Office of Artificial Intelligence, 2022, online (gov.uk).

^{52 &}quot;Establishing a pro-innovation approach to regulating AI," gov.uk.

^{53 &}quot;Establishing a pro-innovation approach to regulating AI," gov.uk.

⁵⁴ "AI Standards Hub," *Alan Turing Institute*, 2022, online (<u>aistandardshub.org</u>).

⁵⁵ Dennehy, Fiona. "New initiative on standards for artificial intelligence launches in the UK," The Alan Turing Institute, 12 Oct. 2022, online (turing.ac.uk).

⁵⁶ "Al Assurance roadmap," gov.uk.

process standards for AI, such as quality- and risk-management standards, and recommends consolidation around terminology and useful ways of measuring bias. Although standards, the CDEI says, offer a critical common foundation for how society expects AI systems to be developed, used, and overseen, they are insufficient on their own. Rather, it will be vital to provide trustworthy information about whether these standards are being followed. According to the CDEI, that means creating a multi-stage assurance process with impact assessments, and standards that support conformity assessment tools.⁵⁷

To ensure AI systems are safe and otherwise trustworthy, the CDEI notes that many AI systems and organizations will need to be assured against a combination of standards of different types (for example, both management standards and standards for measuring bias in particular contexts). AI risk or impact assessments, certification schemes, and other assurance techniques will be shaped by the emerging constellation of standards being determined in international, regional, and national standards-development organizations.⁵⁸ The following chart (Figure 2) from the CDEI's roadmap illustrates a spectrum of prospective AI assurance tools. It should be noted, however, that certification, performance testing and verification and validation are all forms of Conformity Assessment, and a large part of this is accreditation, which is not reflected in the figure below.

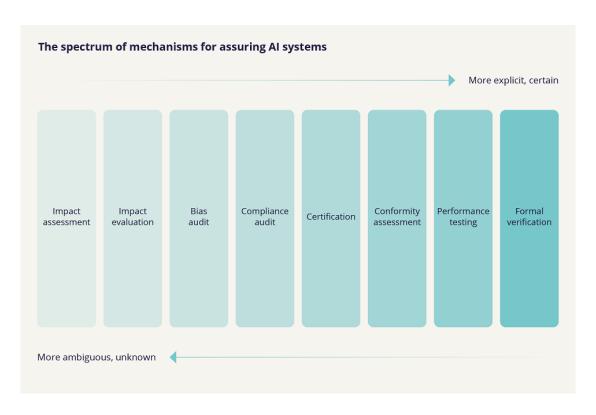


Figure 2: The spectrum of mechanisms for assuring AI systems, ranging from ambiguous and unknown to explicit and certain. Source: Centre for Data Ethics and Innovation.

_

⁵⁷ "Al Assurance roadmap," gov.uk.

⁵⁸ "Human Rights Implications of Algorithmic Impact Assessments," <u>carrcenter.hks.harvard.edu</u>; "The roadmap to an effective AI assurance ecosystem," <u>gov.uk</u>; "A survey of artificial intelligence risk assessment methodologies," <u>trilateralresearch.com</u>.

United States of America

The core of the U.S.'s approach to AI governance stems from a 2019 Executive Order, which requested the development of standards for AI and set out the need to maintain civil liberties, privacy, and American values while continuing to drive AI innovation.⁵⁹

Key tools in the U.S.'s approach to AI include the *National AI Initiative Act (NAIIA)* (2020), the Blueprint for an AI Bill of Rights (2022), the proposed *Platform Accountability and Consumer Transparency Act (PACT)* (2021), and the proposed *Algorithmic Accountability Act (AAA)* (2022). ⁶⁰ The *NAIIA* became law in 2021. It provides for federal coordination on AI research and development but leaves states the power to pass independent legislation. The *NAIIA* also outlines several goals, such as American leadership in AI research and development, including the development and use of trustworthy AI, and preparation of the U.S. workforce for the impact of AI on all sectors of the economy and society. Those portions of the *NAIIA* emphasizing trustworthy AI outline the necessity of a multifaceted approach, naming the development of standards and assessment tools as integral elements of this approach.

While the AAA is not expected to pass, it indicates that U.S. legislators are considering a risk-based approach to regulating AI. It would require companies to conduct impact assessments of AI systems in accordance with regulation set out by the Federal Trade Commission (FTC) and give the FTC authority to conduct bias impact assessments.

The proposed Blueprint for an AI Bill of Rights sets out voluntary guidelines entities can follow to better safeguard against AI risks. These guidelines are based on five overarching principles seemingly informed by the OECD AI Principles, being: notice and explanation, safe and effective systems, non-discriminatory algorithms, data privacy and agency, and options for recourse and opting out of automated systems. ⁶¹

Like the EU's *Digital Services Act*, the proposed *PACT* would impose transparency requirements on social media companies regarding the use of recommender systems. ⁶²

In regard to standardization, the aforementioned 2019 Executive Order asked the National Institute of Standards and Technology (NIST) to create "a plan for Federal engagement in the development of technical standards and related tools in support of reliable, robust, and trustworthy systems that use AI technologies." ⁶³ In response, NIST developed its AI Risk Management Framework, an initial draft of which it released in 2022. ⁶⁴ Soon after, NIST published a paper, "Towards a Standard for Identifying and Managing Bias in Artificial Intelligence," proposing a strategy for managing AI bias and identifying types of bias that occur frequently in the context of AI that can contribute to significant societal harms. ⁶⁵ Previously, in 2021, NIST held a workshop on AI measurement and evaluation and proposed a method to evaluate users' trust in AI. ⁶⁶

⁵⁹ "Maintaining American Leadership in Artificial Intelligence," Federal Register, 14 Feb. 2019, online (federalregister.gov).

⁶⁰ "National Artificial Intelligence Initiative Act of 2020," 116th Congress, 1 Jan. 2020, online (congress.gov); "Algorithmic Accountability Act of 2022," 117th Congress, 3 Feb. 2022, online (congress.gov); "Blueprint for an Al Bill of Rights," Office of Science and Technology Policy, Oct. 2022, online (whitehouse.gov); "Platform Accountability and Consumer Transparency Act," 117th Congress, 17 Mar. 2021, online (congress.gov).

^{61 &}quot;Recommendation of the Council on Artificial Intelligence," Organisation for Economic Co-operation and Development, 21 May 2019, online (oecd.org).

⁶² Recommender systems are Al systems that provide recommendations or suggest relevant items to users, for example, social media algorithms that suggest videos to users based on their usage history. See coons.senate.gov.

⁶³ "Maintaining American Leadership in Artificial Intelligence," <u>federalregister.gov</u>.

⁶⁴ "Al Risk Management Framework: Initial Draft," *National Institute of Standards and Technology*, 17 Mar. 2022, online (nist.gov).

⁶⁵ Schwartz, Reva, et al. "Towards a Standard for Identifying and Managing Bias in Artificial Intelligence," National Institute of Standards and Technology, 15 Mar. 2022, online (nist.gov).

⁶⁶ Stanton, Brian and Theodore Jensen. "Trust and Artificial Intelligence," National Institute of Standards and Technology, 2 Mar. 2021, online (nist.gov).

In the longstanding Circular No. A-119 on federal development and use of voluntary consensus standards and conformity assessment, the U.S. outlines its commitment to using international standards whenever possible. This document also notes the federal prohibition on U.S. government agencies engaging in any standards-related activities that cause unnecessary commerce barriers.⁶⁷

Singapore

Given its size and status as a city-state, Singapore tends to focus on releasing regulation more akin to guidelines that help its businesses and individuals comply with the standards attached to other markets. In that vein, Singapore has launched several large-scale AI policy initiatives. In 2018, it established its Advisory Council on the Ethical Use of AI and Data, the role of which is to advise the government on ethical, policy, and governance issues of AI and to support the government in crafting guidance for private-sector organizations. ⁶⁸

The Smart Nation and Digital Government Office launched the "National Artificial Intelligence Strategy" in 2019. It aims to make Singapore a leader in developing and deploying scalable and impactful AI solutions.⁶⁹

That same year, Singapore released its Model AI Governance Framework. Updated in 2020, the framework provides detailed guidance to private-sector organizations for addressing ethical and governance issues when deploying AI systems.⁷⁰ The Implementation and Self Assessment Guide for Organizations (2020) complements this framework by providing an extensive list of industry examples and practices.⁷¹ Singapore launched sector-specific AI Programmes in 2021, including for government and finance, which are broadly defined initiatives to support AI development in particular industries.⁷²

In 2022, the Infocomm Media Development Authority and Personal Data Protection Commission launched a voluntary AI governance testing framework and toolkit called AI Verify. The program aims to promote transparency between companies and their stakeholders and promote greater public trust in AI by allowing AI developers and owners to verify the claimed approach, use, and performance of their systems against a set of principles through standardized tests. Currently a minimum viable product, the framework is being offered as a pilot for organizations that wish to test it and provide feedback before its next iteration. Outputs of the project are being actively shared with national standards bodies in other countries.

China

Key Al initiatives out of China include its New Generation Artificial Intelligence Plan (2017), Ethical Norms for New Generation Artificial Intelligence (2021), Artificial Intelligence Standardization White Paper (2021), National Standardization Development Outline (2021), and the Internet Information Service Algorithmic Recommendation Management Provisions (IISARMP) (2022).⁷⁴

⁶⁷ "Circular No. A-119 Revised," Office of Management and Budget, 10 Feb. 1998, online (whitehouse.gov).

⁶⁸ "Singapore's Approach to AI Governance," Personal Data Protection Commission Singapore, online (pdpc.gov).

⁶⁹ "National AI Strategy," Smart Nation Singapore, Nov. 2019, online (smartnation.gov).

⁷⁰ "Model Artificial Intelligence Governance Framework Second Edition," *Infocomm Media Development Authority, Personal Data Protection Commission Singapore*, 21 Jan. 2020, online (pdpc.gov.sg).

^{71 &}quot;Singapore's Approach to AI Governance," <u>pdpc.gov</u>.

⁷² "Two New National Al Programmes Launched," Smart Nation Singapore, 8 Nov. 2021, online (smartnation.gov).

⁷³ "Singapore's Approach to Al Governance," <u>pdpc.gov</u>.

⁷⁴ "Full Translation: China's 'New Generation Artificial Intelligence Development Plan,'" *DigiChina Stanford University*, 1 Aug. 2017, online (digichina.stanford.edu); "Translation: Ethical Norms for New Generation Artificial Intelligence Released," *Center for Security and Emerging Technology*, 21 Oct. 2021, online (cset.georgetown.edu); "The Chinese Communist Party Central Committee and the State Council Publish the "National Standardization"

Standards feature prominently in Beijing's plans for Al governance, both domestically and internationally. The main standards body for Al in China is the Standardization Administration of China (SAC). The SAC is China's official representative to the ISO/IEC and manages China's National Committee to the International Electrotechnical Commission (IEC) and other international and regional fora. The SAC developed China Standards 2035, an overarching blueprint articulating the country's plans to lead international standards-making in key emerging technologies.⁷⁵

The 2021 Ethical Norms sets out goals for AI management-system standards, including: actively integrating ethical standards into AI-management processes and ensuring the sustainability of AI, clarifying the responsibilities of different stakeholders, and improving risk assessment in AI development to strengthen risk prevention. This ethical overlay is also present in the IISARMP; Article 4 mandates algorithmic recommendation systems to abide by social, commercial, and professional ethics and "respect the principles of fairness and justice, openness and transparency, science and reason, and sincerity and trustworthiness."⁷⁶

The 2021 National Standardization Development Outline reiterates Beijing's investment in AI standards and conformity assessment. It delegates greater responsibility to industry, lays out standards for AI development and deployment, and aligns these standards with international ones. 77 Also on the international front, China has influenced standards in countries along the Belt and Road Initiative through bilateral agreements. 78

China's Ministry of Industry and Information Technology (MIIT) and Ministry of Science and Technology (MOST) are also actively involved in AI and standardization work. Relevant bodies under MIIT include the National Artificial Intelligence Standardization General Working Group, the Artificial Intelligence Subcommittee, the China Electronics Standardization Institute, and the China Academy of Information and Communications Technology. The National Governance Committee for the 2017 New Generation Artificial Intelligence Plan sits under MOST.

Development Outline"," Center for Security and Emerging Technology, 19 Nov. 2021, online (cset.georgetown.edu); "Artificial Intelligence Standardization White Paper, 2021 Edition," Center for Security and Emerging Technology, 21 Oct. 2021, online (cset.georgetown.edu); "Internet Information Service Algorithmic Recommendation Management Provisions," DigiChina Stanford University, 10 Jan. 2022, online (digichina.stanford.edu).

⁷⁵ Rühlig, Tim Nicholas. "Technical standardisation, China and the future international order: A European perspective," *Heinrich Böll Stiftung*, Feb. 2020, online (<u>ui.se</u>); Gargeyas, Arjun. "China's 'Standards 2035' Project Could Result in a Technological Cold War," *The Diplomat*, 18 September 2021, online (<u>thediplomat.com</u>).

^{76 &}quot;China's Regulation of Internet Recommender Systems: What U.S. Companies Should Know," JD Supra, 1 Mar. 2022, online (idsupra.com).

^{77 &}quot;The Chinese Communist Party Central Committee and the State Council Publish the "National Standardization Development Outline," *Xinhua News Agency*, 19 Nov. 2021, online (cset.georgetown.edu); Sheehan, Matt, Marjory Blumenthal, and Michael R. Nelson. "Three Takeaways From China's New Standards Strategy" 28 Oct. 2021, online (carnegieendowment.org).

^{78 &}quot;Assessing China's Digital Silk Road Initiative," *Council on Foreign Relations*, online (<u>cfr.org</u>); He, Alex. "The Digital Silk Road and China's Influence on Standard Setting," *Center for International Governance Innovation*, 4 Apr. 2022, online (<u>cigionline.org</u>).

Section II: International standards organizations

The International Organization for Standardization and the International Electrotechnical Commission

In 2017, the ISO and the IEC created a joint technical committee on AI: ISO/IEC JTC 1/SC 42 (SC 42).⁷⁹ At the time of writing, this committee has published 16 international standards for AI, with an additional 25 under development. Of these published standards, most appear to be technical reports. In other words, they read like high-level background documents, and by design do not address granular issues or normative content. As discussed below, it will be vital for the Collaborative to properly contextualize such standards into specific, actionable guidelines that can be understood and implemented by organizations within a given industry.

A few standards currently under development are worthy of specific mention. First is the aforementioned AIMS standard, which was successfully proposed by the Canadian mirror committee to ISO/IEC JTC1 SC 42.80 Its goal is to enable organizations to demonstrate the implementation and continual improvement of processes and indicators for developing and using AI, such as bias, fairness, inclusivity, safety, security, privacy, accountability, explainability, and transparency.81 The AIMS standard forms the basis of the AIMS pilot.

Another standard under development at ISO/IEC JTC1 SC 42of note is one for AI impact assessments (ISO/IEC 42005). 82 Given the emerging role of risk and impact assessments in assuring AI systems and products and in complying with AI regulations in particular, this standard is likely to take on an important and practical role in the near term.

A recent German proposal to ISO/IEC JTC1 SC 42 entitled "Information Technology—Artificial Intelligence—Requirements for bodies providing audit and certification of artificial intelligence management systems" is also worth noting due to its likely relevance to Al assurance ecosystems.

Institute of Electrical and Electronics Engineers Standards Association

The Institute of Electrical and Electronic Engineers Standards Association (IEEE SA) is a not-for-profit organization focused on developing international standards to advance technology and technological innovation. The IEEE SA has developed a series of AI standards: Ethics in Action in Autonomous and Intelligent Systems. ⁸³ Standards from this series include those relating to impact assessments of AI on human wellbeing, the transparency of autonomous systems, addressing ethical concerns, data privacy, child and student data governance, virtual classroom security, employer data governance, ethically driven robotics, environmental social governance, trustworthiness of news sources, personal privacy terms, emulated empathy, AI literacy, and algorithmic bias. ⁸⁴

^{79 &}quot;ISO/IEC JTC 1/SC 42 Artificial Intelligence," International Organization for Standards, online (iso.org).

^{80 &}quot;ISO/IEC DIS 42001," <u>iso.org</u>.

^{81 &}quot;Canada 'AIMS' to raise the bar for AI development and use through standardization," Standards Council of Canada, 10 Nov. 2020, online (scc50ccn.ca).

^{82 &}quot;ISO/IEC 42005 Information technology — Artificial intelligence — Al system impact assessment," International Organization for Standards, online (iso.org).

^{83 &}quot;Ethics in Action in Autonomous and Intelligent Systems," Institute of Electrical and Electronics Engineers, online (ieee.org).

^{84 &}quot;Ethics in Action," ieee.org.

Summaries of pertinent standards

Organization	Standard	Description		
ISO/IEC	ISO/IEC 23894 Information technology — Artificial intelligence — Guidance on risk management	This standard provides guidance on managing risks when developing and deploying AI systems. It also can assist organizations to integrate risk management into AI activities. Based on general risk-management standards, the document also notes idiosyncrasies of AI and concomitant considerations, for example, diversity, explainability, and transparency.		
	ISO/IEC 42005 Information technology — Artificial intelligence — AI system impact assessment	This standard contains guidelines for performing impact assessments on AI systems, including when and how to both perform such assessments and complete adequate documentation. It also explains how to integrate impact assessments into AI risk management processes and AIMS.		
	ISO/IEC 38507:2022 Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations	This standard targets the governing bodies of organizations, requiring them to implement policies to ensure appropriate and responsible AI use via clearly defined and agreed-upon delegation processes and chains of responsibility, accountability, and authority. It encourages compliance processes that reflect the speed, scope, and sophistication of AI systems.		
	ISO/IEC 42001 Information Technology — Artificial intelligence — Management system	As discussed above, this standard takes a risk-based approach and targets AIMS, outlining guidelines for measuring effectiveness and efficiency of these systems, as well as for the responsible development and use of such systems that meet applicable regulatory requirements. It is designed to be auditable and certifiable.		
IEEE SA	The IEEE Applied Artificial Intelligence Systems (AIS) Risk and Impact Framework Initiative	Similar to the ISO/IEC document, this initiative aims to propose a risk assessment and mitigation paradigm based on previous models but tailored to AI.		
	IEEE 7010-2020: IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being	This recommended practice provides measures of wellbeing with the goal of creating a wellbeing impact assessment that identifies impacts on and safeguards and enhances human wellbeing throughout the life cycle of autonomous systems in various sectors.		
	IEEE P2863: Recommended Practice for Organizational Governance of Artificial Intelligence	Similar to the ISO/IEC standard, this recommended practice outlines criteria for trustworthy AI, such as transparency, accountability, and safety, as well as guidance on how to responsibly develop or use AI, such as auditing, training, and complying with regulations.		
	IEEE 7000-2021: IEEE Standard Model Process for Addressing Ethical Concerns during System Design	This standard establishes processes to consider ethical concerns during concept exploration and development, including stakeholder consultation and the traceability of ethical values through design, operations, and organizational policies and procedures. The IEEE SA has also developed an ethics certification program for AI. ⁸⁵		

Table 2: Summary of key international standards on Al.

⁸⁵ "The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS)," *Institute of Electrical and Electronic Engineers Standards Association*, online (<u>standards.ieee.org</u>).

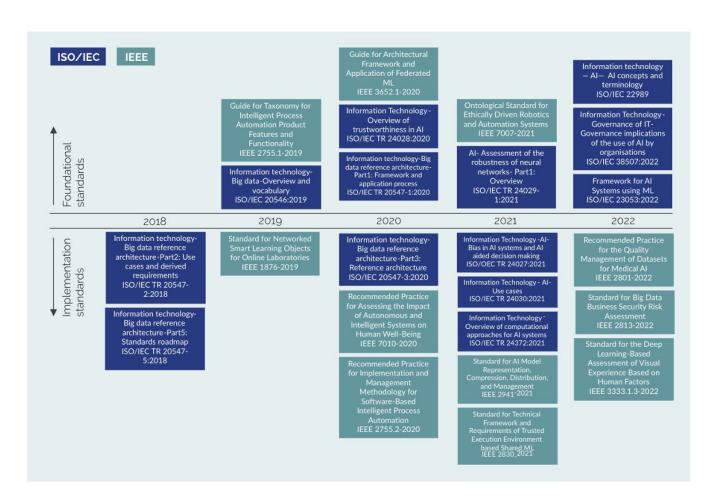


Figure 3: Al governance standards published by ISO/IEC and IEEE from 2018 to 2022. ⁸⁶ Source: University College London.

⁸⁶ Alania, Ana, et al. "Looking Ahead: The Role of Standards in the Future of Artificial Intelligence (Al) Governance," *University College London*, Aug. 2022, online (<u>ucl.ac.uk</u>), p. 16.

PART II: ANALYSIS

Section I: Introduction

States, regions, and international bodies around the world are signaling a reliance on technical standards as one of the principal and common methods for governing AI, even if it is unlikely that technical standards will emerge as a silver bullet for international AI cohesion and interoperability. Additionally, the current international regulatory landscape is increasingly divergent. As such, given the sheer number of bodies turning to technical standards, along with the potential fragmentation in standards and its likely consequences, Canada must better understand, help develop, and promote the development and adoption of cohesive, international AI standards to ensure international interoperability and market access for Canadian AI organizations.

Furthermore, although broad international standards will form an important part of the AI regulatory landscape, these instruments can only become actionable and useful to individual players if efforts are made to help businesses contextualize standards in particular industries and contexts through conformity assessment tools. While this work will not happen at the ISO/IEC, it is important to acknowledge the limitations of horizontal (general, inter-sectoral) standards and the importance of the emergence of vertical (sector- or application-specific) standards. The AIMS pilot is an example of the type of effort the Collaborative could continue to pursue, with expanded capacity, to contextualize and render useful international standards, while developing guidelines on implementing or conforming with these standards. In the case of the AIMS pilot, work is currently focused on financial services use cases, suggesting that part of the pilot's purpose is to develop sector-specific implementation guidelines.

Section II: Key international actors

Over the past several years, AI and other frontier technologies have taken on increasing geopolitical significance, due as much to their impacts on shared societal norms and values as to their potential to define global economic competitiveness. As the importance of standards and conformity assessment to international trade, investment, competitive advantage, and values has intensified, so too has competition in developing AI standards and conformity assessment, both within international bodies and as a matter of national policy. Since standards and conformity assessment can play key roles in AI governance by supporting regulation, codifying knowledge, enhancing coordination, increasing trust and safety, and ensuring interoperability, they are not only a means of attaining AI leadership in a general sense, but are also an avenue through which countries can operationalize and export values.⁸⁷

Three key players have emerged in this context: the U.S., China, and the EU. All three are strong economic blocs with large markets and varied views on and approaches to AI standardization. The EU was the first to introduce a comprehensive regulatory scheme. When the EU passes the AI Act, companies wishing to place AI products on the EU market will have to comply with its essential requirements and may need to demonstrate that compliance through conformity assessment. Thus, although some commentators have argued that the AI Act is poised to generate only a partial or fragmented "Brussels effect," in which EU standards could become *de facto* international standards, the question of whether European harmonized standards will ultimately differ from international standards remains an important topic to monitor. 88

87 Callegari, Agustina, et al. "Looking ahead: the role of standards in the future of Al governance," British Standards Institute, Aug. 2022, online (ucl.ac.uk).

⁸⁸ Siegmann, Charlotte and Markus Anderljung. "The Brussels Effect and Artificial Intelligence: How EU regulation will impact the global AI market," arXiv, 23 Aug. 2022, online (arxiv.org).

The U.S. is a global leader on AI talent, research and development, and commercialization, and has begun to develop initiatives for standardization.⁸⁹ It has also identified standards as an important area for international cooperation among its allies—both for economic competitiveness and the protection of shared norms—and has undertaken a variety of initiatives to work with its allies on developing harmonized international standards, at least in part to counter China's influence in global technology markets.

China has also launched several initiatives on AI standards and conformity assessment both nationally and internationally, as outlined above. It has a robust industrial policy that prioritizes the development of national standards, investment in international digital infrastructure projects that incentive their adoption, and participation in international standards bodies. The influence of China on international technology standards, and the combination of the "push" and "pull" effects of its policies, has been described as the "Beijing Effect." ⁹⁰

All three of these key players have large markets and a great deal of international influence, and too much divergence in their approaches to AI regulation and standardization could lead to significant fragmentation of the international AI landscape. Against this backdrop, international organizations and both bilateral and multilateral organizations such as the G7, the U.S.-EU Trade and Technology Council, and the Global Partnership on AI have emerged as prominent for a for promoting alignment on AI standards and conformity assessment.

Section III: The role of standards and conformity assessment

Despite vastly differing approaches to AI regulation, each of the regions highlighted in this paper has indicated a clear intent to use standards, conformity assessment, and other assurance tools, increasing the importance of such instruments. There are early signals to indicate the EU will pursue an assurance ecosystem, while also implementing the AI Act, a strict, risk-based, top-down framework. Thus, the EU is likely to require organizations developing or using AI to implement many processes to comply with the manifold requirements, including conformity assessment. This approach may be expensive and difficult, requiring quite significant effort from organizations and the public sector. Further, while this seems to suggest AI systems in the EU will be highly regulated, it is not clear these detailed and onerous requirements will translate into effective regulation.

By contrast, the UK has signaled a vastly different approach: one that is voluntary, principles-based, and context-specific, rather than strictly regulating responsibilities and requirements. While there is no national regulation dictating requirements for AI, mandatory sectoral regulation could still emerge. This approach eschews a common regulatory framework to address general, broadly applicable risks. The UK has made it clear that innovation and economic growth are priorities when it comes to regulating AI, demonstrating a clear reluctance to formally regulate AI at the broader state level. Under this approach, AI regulation, including any mandatory measures, is largely left to existing regulators in specific sectors. While the EU plans to create new rules and regulatory agencies and the UK is aiming to avoid this approach, both will need some type of assurance ecosystem. Indeed, the UK's assurance roadmap signals an intention to do so, despite some mixed signals.

Singapore is advanced when it comes to implementing practical standardization and governance solutions on a national level, a choice perhaps influenced by its status as a city-state. As mentioned above, however, this status means that it will need to adapt to international AI standards to remain competitive on the global stage. As such, Singapore does not appear to be planning to pursue a prescriptive approach either, opting instead for the creation of a trusted ecosystem that balances

20

^{89 &}quot;The Global AI Index," tortoisemedia.com.

⁹⁰ "The Digital Silk Road and China's Influence on Standard Setting," <u>cigionline.org</u>; Erie, Matthew S. and Thomas Streinz. "The Beijing Effect: China's "Digital Silk Road" as Transnational Data Governance," *New York University Journal of International Law and Politics*, 2021, online (<u>nyujilp.org</u>).

regulatory guidance with innovation and adoption. Singapore will likely continue to introduce guidelines and non-binding frameworks that align with its National AI Strategy and AI Verify, as well as sector-specific AI programs.⁹¹

At present, the U.S. does not seem to have settled on a regulatory approach, aside from an intention to avoid regulation that could restrict innovation. Its tools include frameworks for risk management, bias, and trust, as well as voluntary principles. Risk has begun to become a more prominent topic of discussion, which some observers believe may suggest increasing alignment with the EU approach. However, most initiatives remain focused on avoiding individual harms, such as algorithmic discrimination. Moreover, the U.S. has shown that standards will play an important role in its regulatory efforts and has committed to using international standards whenever possible. Thus, although the U.S. has taken several steps indicating strong consideration of AI policy, including the directives, guiding documents, and international standardization initiatives mentioned above, there is a lack of a clear, nation-wide approach to regulating AI. This may be an opportunity for alignment on the issue with different world leaders. Given the various signals from the U.S. indicating that the government may see strict regulation as a barrier to AI innovation, it is possible the U.S. will align closely with the approach currently being signalled by the UK.

China's approach to AI regulation primarily is a top-down, prescriptive, rules-based framework, supported by other approaches including the conformity assessment tools and AI ethical principles. The Cyberspace Administration of China is the key regulator behind the country's highly prescriptive approach, which includes a draft set of thirty rules for recommender algorithms, a three-year roadmap for the governance of all algorithms, and IISARMP. ⁹⁴ Chinese actors will be expected to conform to specific requirements regarding algorithmic transparency and explainability, as well as the dissemination of information. Standards also feature prominently in Beijing's plans for AI governance, both domestically and internationally.

Despite these differing regulatory approaches, signals from each of these regions indicate strong intent to use and recognition of the importance of various assurance techniques. The *AI Act* includes mandatory conformity assessment for high-risk AI. The CDEI in the UK has centred the use of standards along with impact assessments and conformity assessment. The U.S. has signalled strong support for standards in several initiatives, including the *NAIIA*'s view of standards as integral; NIST rhetoric with an emphasis on standards dealing with risk, bias, and consumer trust; and Circular No. A-119, which outlines a commitment to international standards.

Moreover, the U.S., the EU, and the UK have each put forth a regulatory framework that uses impact assessments, despite differences in the overall frameworks. Although each jurisdiction proposes to use impact assessments in a different way, it will nevertheless be important to develop a standard impact assessment methodology. The complications that would arise if AI systems were classified as high risk in some jurisdictions, and low risk in others, is clear.

As mentioned above, ISO and the IEC have already identified the importance of a standard international impact assessment methodology. A work item proposal recently accepted by the two bodies aims to standardize the development process and structure of AI impact assessments.⁹⁵

Section IV: Consequences of a fragmented international landscape

⁹¹ This would be in line with Singapore's typical actions, given its position as a city-state and desire to retain access to international markets.

⁹² Engler, Alex. "The EU and U.S. are starting to align on AI regulation," Brookings, 1 Feb. 2022, online (brookings.edu).

^{93 &}quot;Circular No. A-119 Revised," whitehouse.gov.

⁹⁴ "IISARMP" (digichina.stanford.edu); Goldkorn, Jeremy. "Government announces three-year plan to tame China's algorithms," *The China Project*, 29 Sept. 2021, online (thechinaproject.com).

⁹⁵ ISO/IEC 42005, (iso.org).

Despite all of the above—signals, published instruments, or membership in international standardization bodies—there is no guarantee every country will follow internationally recognized standards or principles. Each of the major players highlighted above has the power to export its regulations and standards, and thus values, through a first-mover advantage and consequently shape the international standards landscape. Should this occur, it could spell trouble for international cohesion and complicate international interoperability and market access. Further, allowing any particular national or regional standards to become *de facto* international standards would not only fragment the landscape, but also take away from the legitimacy and power of international standards created through more democratic processes. The widespread use of international standards is also important because all participating members of ISO/IEC JTC1 SC42, including Canada, are given the opportunity to provide input into these standards and participate in their development.

Take the EU as an example. Many consider it to be a jurisdiction that may develop and export standards that differ from existing international standards. This is illustrated by the *AI Act*, which references the importance of "harmonized standards" and delegates CEN and CENELEC to develop standards for AI. 96 Should the EU adopt harmonized standards different from those of international bodies, this could not only create barriers for Canadian companies to access its market, but the Brussels effect could lead to European standards turning into *de facto* global standards, if adopted by a critical mass of smaller states looking to retain EU market access. Such a scenario would cause issues for interoperability and market access by forcing states to choose between different standardization regimes, creating large hurdles for organizations attempting to access markets implementing another set of standards. Countries with smaller domestic markets, like Canada, would likely bear the brunt of such a fracture. To illustrate, under a split regime, a Canadian AI start-up offering a product that determined whether a bank should approve a business loan may be required to create two different versions of that product if it wished to retain global market access: one compliant with EU standards and one compliant with international standards. This would be cost prohibitive for many Canadian businesses and, in some cases, practically infeasible, should there be hard incompatibilities between the two standards. Similarly, testing a product for compliance with two or more different regimes could also involve prohibitive costs.

That said, European harmonized standards are unlikely to differ substantially from those already published by ISO/IEC or IEEE SA. While it is possible the EU could inject some differences in its standards to reflect its priorities and the essential requirements of the *AI Act*, it is unlikely to be beneficial for the bloc to differ from internationally recognized standards that it was involved in developing, particularly given the significant resources required to develop new standards. Moreover, the 1991 *Vienna Agreement* establishes a process by which standards developed by either ISO or CEN are notified for approval by the other (this agreement being similar to the IEC/CENELEC Dresden Agreement from 1996). It sets out the primacy of international standards, while acknowledging that the EU market may require certain standards that have not yet been developed at the international level. As such, European standards organizations are expected to develop entirely new standards only when an international standard does not already exist. While European values may influence some of the harmonized standards that will be developed under the *AI Act*, these are likely to be granular, sector-specific standards that do not have a significant impact on the international standards landscape.

It is important to look to such examples to illustrate the consequences that could arise from a fragmented international landscape. As addressed above, there is major divergence in regulatory approaches to AI around the world. However, by looking to common international standards for AI, and improving cohesion through the use of tools such as conformity

⁹⁶ A European standard developed by a recognised European Standards Organisation: CEN, CENELEC, or ETSI. It is created following a request from the European Commission to one of these organisations. Manufacturers, other economic operators, or conformity assessment bodies can use harmonised standards to demonstrate that products, services, or processes comply with relevant EU legislation. "Harmonized Standards," *European Commission*, no date, online (<u>ec.europa.eu</u>).

⁹⁷ "Agreement on technical cooperation between ISO and CEN (Vienna Agreement)," International Organization for Standardization, European Committee for Standardization, June 1991, online (boss.cen.eu).

^{98 &}quot;Vienna Agreement," boss.cen.eu.

assessment and impact assessments, these risks can be mitigated, even if technical solution to AI governance.	standards alone may not be a complete

PART III: RECOMMENDATIONS

Section I: Introduction

In this complex and fragmented landscape, it will be critical for Canada to identify international and regional partnerships that could support Al-standards initiatives as part of a greater collaboration on Al governance. International high-level initiatives align states on broad principles and policy directions, but more targeted bilateral agreements could provide political capital and opportunities for Canadian companies and experts to contribute to standards development organizations.

Canada can also advance its standardization objectives by investing in targeted initiatives in particular sectors. This could occur, for instance, through sandboxes or pilots aimed at developing adoption or implementation guidance for SMEs in certain industries, conformity assessment schemes, approaches, or programs for standards incorporated by reference in legislation, or support for Canadian companies developing technologies that inform the development or operationalize the adoption of AI standards and conformity assessment. The Collaborative could play an important role in defining a methodology for selecting and prioritizing such initiatives.

Finally, Canada could undertake several initiatives to support the development or adoption of standards by meaningfully consulting a diverse range of stakeholders for input. For example, in developing a strategy for AI standardization in Canada, the Collaborative should pay particular attention to Indigenous perspectives on AI development and adoption and the realities of Canada's technology ecosystem (which consists mainly of start-ups and SMEs) while encouraging participation by a variety of civil-society actors in standards development.

The following section will offer recommendations for the Collaborative's consideration by building upon the identified issues and opportunities.

Section II: General recommendations

Recommendation I: Identify priority standards and conformity assessment schemes, approaches and programs

An overview of the international AI regulation and standardization landscape reveals many governments allocating particular focus to risk management, bias, quality management, governance, and impact assessments. As the above recommendations discuss, these focus areas provide arenas for increased bilateral partnerships, including opportunities to collaborate with NIST on the development of risk-management (and perhaps bias-management) frameworks, and collaboration with European standardization organizations on developing joint conformity assessment schemes, approaches and programs.

Beyond bilateral partnership, however, it will be important for the Collaborative to recognize the standardization areas for which there has been a great deal of convergence. These areas—risk management, bias, quality management, governance, and impact assessments—present a helpful initial list from which the Collaborative can work to identify standardization priorities. Taking this list as a starting point will aid the Collaborative in identifying priorities for which there are greater opportunities for bilateral and international alignment and collaboration.

Promoting the development and adoption of these areas that have been globally recognized as important will be integral to maintaining cohesion with international frameworks and markets. For example, the Collaborative may consider identifying priorities for businesses in areas where industry is struggling with risk management. Solutions to these priority areas could

be discussed as a working group, piloted, and launched in regulatory sandboxes to allow the Collaborative to both inform risk-management standards and help incentivize their adoption in Canadian markets.

It is also worth noting that some of the above-mentioned priorities fall into organizational-process standards (quality management, risk management, governance) while others are more product-centric. Because of this, there has been recent discussion amongst conformity assessment experts about the possibility of merging organizational-process standards and product-centric standards in a joint conformity assessment program. This possibility is discussed in more detail below.

Recommendation II: Enhance Canadian participation in developing international standards

Canada should aim to punch above its weight when it comes to influencing the quality, pervasiveness, and speed of development of international standards. In fact, given the small size of the Canadian domestic market, influencing international standards and thus securing access to international markets for domestic organizations will be necessary to ensure the survival of the Canadian Al economy.

There are opportunities for Canada to play a key role in not only developing international standards through ISO and IEC mirror committees, but also to dig deeper into questions about improved synchronization between Canada's regulatory approach to AI and the standards currently under development at international standards development organizations. For example, are there portions of Bill C-27 that should be better reflected in ISO/IEC JTC1 SC 42documents? How well do definitions and terminology from Bill C-27 map onto the standard on AI concepts and terminology?⁹⁹ At present, Bill C-27 omits certain concepts included in the standard, such as "human-defined objectives."

Further, Canada must encourage its citizens to participate in ISO/IEC mirror committee work, which, in turn, will help to maintain and expand Canada's influence in standards-setting. This includes enhancing the skills of and incentivizing an interdisciplinary group of experts to participate in the ISO/IEC JTC1 SC 42 working groups that have an impact on their interests. Canada needs experts to deliver on certain fundamental elements of standardization, including drafting new work item proposals and participating as project editors or convenors in technical committees. The Collaborative should therefore identify gaps in Canadian participation on priority in ISO/IEC JTC1 SC 42initiatives and develop strategies to increase Canadian membership and participation. These strategies could include early career curricula, participation models, and concerted outreach to existing experts. Existing gaps the Collaborative may wish to explore further include a lack of consumer representation and low levels of membership from academia.

Canada must also continue offering strong support for the development of cohesive international standards. This is vital for ensuring general interoperability, accessibility, and market access, and even more so for stakeholders for whom access and participation is particularly challenging. As Canada pushes for increased mirror committee participation, taking the interests of these stakeholders into account—and developing strategies to enable their participation despite capacity limitations—should remain top of mind.

Finally, given that the creation of a standard impact-assessment methodology will likely have far-reaching impacts on international cooperation, Canada should prioritize participation on the standard for AI impact assessments currently under development in ISO/IEC JTC1 SC42. 100

⁹⁹ ISO/IEC 22989:2022 (iso.org).

¹⁰⁰ That is, ISO/IEC 42005 (<u>iso.org</u>).

Recommendation III: Enhance SME engagement in developing international standards

It is necessary to engage SMEs in AI standardization, from participating in standards development to accessing the resources necessary to successfully undertake conformity assessment against those standards. However, given the limited capacity of SMEs in terms of both funding and employees, achieving such engagement may pose difficulties.

The barriers preventing greater engagement on AI standardization initiatives by SMEs have little to do with a lack of interest or prioritization. SMEs are generally aware they are welcome to participate in standards development and the development of these standards will directly affect their business interests. Yet, most SMEs cannot afford to direct their resources towards anything but the most vital tasks.

To enhance SME engagement, the government could provide funding for additional resources to help SME's participate in standardization, for example, to help gather insights from Canadian SMEs and industry associations, support attendance at ISO/IEC mirror committee meetings, and make contributions on behalf of SME interests. Although this strategy has flaws for example, SMEs will have competing interests—it could offer an effective first step toward increasing accessibility of participation in these processes. Smaller organizations that cannot afford to send an employee to multiple meetings in order to contribute to standards development could communicate their needs to for, example, a SME interest group or dedicated SME liaison. The decreased time and resource commitment required to voice concerns to a dedicated representative would significantly lower barriers to participation and help bring the voice of SMEs to the standardization table. Additionally, funding alternative initiatives that help bolster SME participation and input could be explored.

Recommendation IV: Meaningfully engage with Indigenous peoples

It will be important for the Collaborative to understand the issues and topics most relevant to Indigenous peoples, acknowledging that Indigenous governments have their own perspectives and objectives on these areas. ¹⁰¹ For the Canadian government, AI standards and conformity assessment are a clear priority. As discussed, standardization will be vital to ensure interoperability of AI systems across international borders, market access for Canadian AI firms, continued AI innovation, and more.

However, from a First Nations perspective, although AI standards and conformity assessment are by no means unimportant, this work is not particularly meaningful until issues of Indigenous sovereignty, and particularly data sovereignty, are resolved. Indigenous data sovereignty has been previously raised by the DGSC, and can be defined as "the right of an Indigenous governing body to govern the collection, ownership, dissemination and application of its own data about its communities, members, lands and resources." 102 Often, when the issue of data sovereignty is raised in rooms focused on AI standards and conformity assessment, non-Indigenous stakeholders in the room offer hopes that data sovereignty could be achieved, but generally do not engage on the issue. Yet, for First Nations stakeholders, the issue of sovereignty is indelibly tied to the issue of AI standardization, as attempts to increase trustworthiness and reduce harms through standards should look not only to the development and deployment phases of AI, but also at the collection and use of data to train AI. If the issue of Indigenous data sovereignty is not meaningfully addressed, then standards attempting to mitigate harms to Indigenous populations at the deployment stage will be ineffectual. 103 It is further important for those participating in standardization discussions to

¹⁰¹ We note there are three distinct Indigenous groups within Canada with potentially varying perspectives on AI standardization—First Nations, Inuit, and Métis—and we were only able to gather the First Nations perspective.

^{102 &}quot;Canadian Data Governance Standardization Roadmap," scc.ca.

¹⁰³ John Borrows and Lisa Austin. "The Digital Charter Implementation Act Ignores Indigenous Data Sovereignty," Schwartz Reisman Institute for Technology and Society, 6 Dec. 2022, online (srinstitute.utoronto.ca).

recognize differences not only from a cultural or linguistic point of view; there must also be recognition that First Nations data is coming from a different jurisdiction than Canada.

Further, the Collaborative should be aware that resourcing is a barrier to participation by First Nations organizations and individuals in standards-setting. Nor is the only issue participation in and submissions at mirror committees. The extensive groundwork and resources required to even follow all of the committees, projects, and various conversations around AI standardization is often already too much; there are too many rooms to be in. Here, the Collaborative may again consider the strategy of securing funding for additional resources to bring First Nations, Inuit, and Métis perspectives to ISO and IEC mirror committees.

The Collaborative should not allow the conversation about AI standards and conformity assessment to become completely dominated by larger users and the considerations prioritized by powerful multinational corporations, such as interoperability or innovation. It is vitally important that AI systems are at a scale that is human, manageable, and accountable, and the most valuable AI systems will be those that can have some application to decolonization or stewardship over waters and lands for current and future generations.

It is important for the Collaborative to acknowledge Indigenous sovereignty and work towards the reconciliation of these diverging objectives in much the same way as global governments find avenues toward continued partnership and harmonization. The ultimate goals of consultation on AI standards and conformity assessment must be a strong commitment to reconciliation and a desire to collaboratively design a system that meets the respective needs of both parties. It cannot be a box-checking exercise.

To inform its approach to this recommendation, the Collaborative should consider reaching out to Indigenous groups in Canada first, and then in other countries such as the U.S., Australia, and New Zealand for their perspectives and insights into how they are interacting with their national governments. In Canada, the Collaborative will likely need to consult with Indigenous Services Canada; Crown-Indigenous Relations and Northern Affairs Canada; and Innovation, Science and Economic Development Canada to ensure they are operating in line with Government of Canada best practices. Ensuring that Indigenous sovereignty is acknowledged and that the perspectives of Indigenous stakeholders are brought to international standards development processes may be then another strong area for bilateral or multilateral cooperation, if this resonates with the Indigenous stakeholders in these countries.

Recommendation V: Encourage industry engagement in standards development

Often, industry has more resources and larger budgets than the public sector. In the context of AI governance, there is significant investment from industry players into the discovery and implementation of best practices for AI regulation and responsible AI. Regulators must figure out how to use these resources to ensure favourable outcomes for everyone, from public consumers to academics to SMEs that require support to understand and comply with best practices.

Here, the Collaborative should consider the critical role the SCC could play in bridging the gap between industry players, academics, and other stakeholders by creating new ways for these actors to bring the results of their efforts to committees without necessarily needing to sit in committee seats. Leveraging the strengths of industry and, in particular, its financial power, and those of researchers, such as deep knowledge and a willingness to participate in committees and submit proposals, would be an effective way to increase Canadian participation in AI standards and conformity assessment development and interdisciplinary cohesion and communication. This could also be helpful for regulators, possibly increasing usage of international standards and cohesion between international and Canadian standards.

More specifically, the Collaborative could facilitate collaboration between National AI Institutes under PCAIS and ISO/IEC JTC1 SC 42mirror committees. One potential pathway for doing so could be rendering research and development spending tied to a proposal at a standards committee eligible for tax credits. Structures that promote the translation of research and development initiatives into tangible proposals at standardization committees could increase cohesion between academia, industry, and standards development organizations. Bridging the current gap between these players could bring about transformative breakthroughs, fuelled by pooling the strengths of each.

Recommendation VI: Enhance openness, accessibility, and diversity in standards development

It will be vital for the SCC and the Collaborative to put continual emphasis on enhancing openness, accessibility, and diversity in standards development. It is well known that there are significant barriers to participation in standardization processes, some of which have already been discussed at length in this paper. However, additional concerns about openness and diversity in standards development centre around the fact that often, it is large multinational corporations that have the time, resources, and expertise required to truly shape standards, leading to the development of standards that have been primarily shaped by a fairly homogenous group of actors. Consider, for example, the existence of "pay-to-play" models within some standards development organizations, which provide more votes to members that pay higher subscription fees. Situations such as this, as well as the key role standards will likely play in shaping AI regulation, can lead to regulatory capture. Such a situation is highly undesirable, particularly from a diversity standpoint.

Given the current state of play in standards development, it will be vital for the Collaborative to take a proactive approach toward addressing these existing issues, and promoting the adoption of truly open and democratic standardization processes. Specific recommendations set out above address the importance of promoting SME engagement and respecting Indigenous sovereignty during the standards development process, but it is imperative that the Collaborative remember that the barriers highlighted in these sections exist for numerous other civil-society groups. Ensuring openness, accessibility, and diversity in the standards development process for AI will require recognizing the many different groups that may face barriers to participating in standardization work and actively working to enable their participation.

Recommendation VII: Look globally for inspiration on innovative standards initiatives to support their adoption

Around the world, governments and other bodies have proposed or implemented various strategies to develop or clarify standards for AI. The Collaborative should look at these initiatives, consider if they could bolster Canada's approach to AI standards and conformity assessment, and, if so, spearhead the implementation of similar initiatives in Canada. Four examples of such strategies are the UK's AI Standards Hub, Singapore's governance-testing framework (AI Verify), the Spanish regulatory sandboxes, and the NIST-OECD Catalogue of AI Tools & Metrics.

UK: AI Standards Hub

The UK's AI Standards Hub—a single, central, and comprehensive resource for the British AI community to come together and learn about all pertinent standards—can offer inspiration to the Collaborative in its mission to help drive the development and adoption of standards for AI. A major difficulty in AI standardization is the current level of fragmentation across many actors, including standardization bodies, international and national authorities, and academic or research institutions. This fragmentation, combined with a lack of communication and knowledge about who is working on AI standardization issues, has led to the duplication of work. Creating a single Canadian "hub" for AI standards and conformity assessment would provide education, community, and certainty for stakeholders as they attempt to determine and fulfill obligations related to AI. The Collaborative could temporarily act as Canada's functional equivalent to the AI Standards Hub. However, there may be a need for a more formal and permanent structure that can continue to provide support, education, relevant repositories, and more.

Singapore: AI Verify

Canada could also consider emulating Singapore's AI Verify—the voluntary framework that tests and validates claims made by the owners and developers of AI systems about the approach, use, and performance of their products—to promote transparency and trust between organizations, stakeholders, and the public. Such a tool would again help incentivize the development and adoption of AI standards and conformity assessment.

Spain: Regulatory sandbox

The Collaborative could also consider adopting the use of instruments such as minimum viable products, regulatory sandboxes, and pilot programs. For example, Singapore's AI Verify is currently a minimum viable product, whereby organizations are able to try the framework before it is finalized and provide feedback on future versions.

A regulatory sandbox is a secure environment in which organizations can deploy AI systems, with safety measures in place, to learn what challenges or consequences may result from full deployment of the system. It is an ideal testing environment for projects without a home under current regulatory frameworks. The aforementioned Spanish regulatory sandbox on AI allows innovative AI for the financial-services sector to be tested with real consumers who have provided informed consent in order to determine the safety of new AI and pave the way towards its legal acceptance. Entities apply to enter the sandbox and, if approved, undergo an initial six-month testing period. Certain requirements must be met before testing begins, including the informed consent of participants and certain guarantees and indemnities. ¹⁰⁴

The ubiquity of AI remains a major challenge for regulators, and conducting pilots that provide insight into the specific principles and needs of particular industries or types of organizations alleviates some uncertainty around developing further frameworks for general use. These initiatives engender iterative development processes, give AI organizations a say in the creation of a regulatory product, and generate creative ideas. Industry players can be confident they are operating in a controlled, legal environment. Regulators have the opportunity to observe the deployment and use of new AI products and systems, granting them firsthand insight into how to regulate and mitigate against the harms of such technologies. For researchers, the ability to observe technologies in a controlled environment provides invaluable learning opportunities.

NIST-OECD Catalogue of AI Tools & Metrics

NIST and OECD are currently collaborating on their Catalogue of AI Tools & Metrics. ¹⁰⁵ This constantly evolving catalogue aims to document all of the latest tools and metrics available that can assist AI actors with accountability; implementing AI systems and products that respect human rights; and ensuring those systems and products are fair, transparent, explainable, robust, secure, and safe. Though many such tools and metrics exist, they are typically quite difficult to find, and as NIST and OECD point out, are often absent from AI policy discussions. In this instance, unlike the other examples outlined above, the catalogue is global, meaning there is no need for the Collaborative to consider recreating a similar tool within Canada. Canadian AI actors should be able to locate tools that fit their specific needs, as well as useful metrics, through the existing catalogue. However, it will be very useful for the SCC and the Collaborative simply to be aware of this catalogue as a useful database to use—and perhaps promote—during their work on AI standardization. The Collaborative could consider contributing to such an initiative by developing guidance for the use of appropriate metrics and tools for particular industry use cases.

Section III: Opportunities for bilateral activities

¹⁰⁴ Spanish Regulatory Sandbox (<u>sandboxspain.com</u>).

^{105 &}quot;OECD-NIST Catalogue of AI Tools & Metrics," OECD, no date, online (oecd.ai).

Canada-EU

Collaborate on the expansion of the ISO/IEC AIMS standard into a joint certification scheme

Attention within Canada and at ISO/IEC JTC1 SC42 is currently focused on management-system standards. As noted, Canada successfully introduced the AIMS standard at ISO/IEC, and it is now under development. However, management-system certification alone will not grant organizations access to foreign markets developing their own certification schemes; in some cases, product certification may also be necessary. For example, the *AI Act* seems to require organizations to evaluate AI systems as though they are evaluating a product, compelling businesses to obtain a guarantee, through a CE marking, that the AI system meets European product standards before its deployment. ¹⁰⁶ Yet, organizations may also have to use a management-system evaluation to deliver on certain accountability, monitoring, and quality requirements, thus increasing the burden of compliance.

Both management-system and product certifications may be required by different regimes. As such, it is not a question of whether the *AI Act* or the AIMS standard should be adapted to focus solely on one certification, but rather whether actors should pursue the development of a joint certification scheme in which a single process functions as both a management-system and product certification.¹⁰⁷ The Collaborative can play a part by pushing for an expansion of the ISO/IEC AIMS standard to include this joint certification. In this push, CEN, CENELEC, or ETSI are natural bilateral partners.

Currently, there is no ISO/IEC mechanism to merge certification for products and management systems (even though using both schemes exists in other areas such as food safety). Nor is there a widely recognized product certification for AI yet. The Collaborative is well positioned to urge innovation in this area and push for a joint certification under an accreditation umbrella that would allow Canadian organizations to meet management-system expectations, which will likely be derived from the ISO/IEC AIMS standard, as well as product expectations imposed by the EU. A joint certification like this would widen the market for Canadian participants of all types: not only AI companies, but also organizations that have begun using AI in some capacity. Moreover, joint certification would be another step toward international regulatory interoperability. As such, it is recommended that the Collaborative make efforts towards bilateral partnership with national standardization bodies of EU member states on this initiative but remain open to additional partnerships with other states and regions to further assist in the development and adoption of such a joint certification scheme. Soliciting input from CEN and CENELEC, but also from other national standardization bodies including the British Standards Institute and the American National Standards Institute are recommended to help the Collaborative identify bilateral partnership opportunities to expand the AIMS standard and pilot joint certification.

Canada-US

Conduct pilots on how different businesses implement NIST's AI risk-management framework

There has been much international focus on risk management, and the Collaborative should consider this a priority for AI standardization. Assuming risk management is given priority, a strong opportunity for partnership between the SCC and NIST would be to assist with the latter's AI risk-management framework, specifically by offering to conduct pilots on how businesses across different sectors implement the draft framework. The Collaborative could gather input from these pilots to provide NIST in order to help inform the development of this standard.

¹⁰⁶ "Al Act," <u>eur-lex.europa.eu</u>. Article 16(i).

¹⁰⁷ A "product" in this context typically refers to the end implementation, whether that is in a cloud service, a software program, or delivered through a device, for example.

Work with NIST to map linkages between the draft AI Risk Management Framework and the ISO/IEC JTC1 SC 42 AIMS standard

Another opportunity for bilateral activities with the U.S. may lie in soliciting input from NIST on the ISO/IEC AIMS pilot. Comments and recommendations from NIST would help strengthen this project and may open the door to future collaboration on similar pilots that, as discussed above, will be the type of activities the Collaborative should engage in to develop clearer, more specific guidelines on how to implement or adhere to specific standards for AI.

Section IV: Priority use cases for AI standardization

Selection criteria

Stakeholders have proposed relevant uses of AI to help the SCC and the Collaborative engage with more concrete, granular issues, such as community health data and standardization, digital identity and open banking, and digital food supply chains. However, there is no framework for choosing such examples, which would provide a more robust and rigorous justification for their selection.

We propose the following criteria for identifying priority use cases for standardization:

- I. The use case is the subject of international, national, or local legislation, regulation, or public policy impacting Canadian industry and the public.
- II. The use case involves technical components of trustworthy AI that could help unlock research, innovation, or commercialization in an industry that is important to Canada or the public interest (see also the AI priorities identified by the DGSC). ¹⁰⁸
- III. The use case represents a sector with recognized Canadian leadership.

If each of these criteria, described in more detail below, are met, the SCC and the Collaborative should consider exploring the proposed AI application.

Criterion I: Identified by important legislation, regulation, or public policy

Legislation, regulations, and policies, as evidenced by the discussion of this paper, are being developed rapidly in many jurisdictions. Such proposals operate on all scales, from international to hyperlocal. Many highlight, whether explicitly or implicitly, specific areas of concern for the developers of the policies or their polity. As such, it would be helpful for the SCC and the Collaborative to hone in on these already-identified situations.

For example, in 2021, New York City passed a bill that requires bias audits of automated or AI-assisted employment decision tools. ¹⁰⁹ That same year, the U.S. Equal Employment and Opportunity Commission announced it would review the use of technology in employment decisions. ¹¹⁰ It may thus be worth the SCC and the Collaborative exploring AI systems used in human resources as a way to add value to emerging regulation.

^{108 &}quot;Canadian Data Governance Standardization Roadmap," Standards Council of Canada, 28 June 2021, online (scc.ca).

¹⁰⁹ Lee, Nicole Turner and Lai, Samantha. "Why New York City is cracking down on AI hiring," 20 Dec. 2021, online (brookings.edu).

¹¹⁰ "EEOC Launches Initiative on Artificial Intelligence and Algorithmic Fairness," *U.S. Equal Employment Opportunity Commission*, 28 Oct. 2021, online (eeoc.gov).

As another example, the *AI Act* has identified some high-risk AI applications that could affect the fundamental rights of citizens, for example, education, border control, administration of justice and democratic processes, critical infrastructure, safety components of products, law enforcement, and essential private and public services. As such, the Collaborative could potentially examine the use of AI in these contexts.

Finally, impact assessments and conformity assessment have been identified by several governments and non-governmental organizations in their proposals, strategies, roadmaps, and other documents on AI standardization. It is clear that these assessments will form a significant part of upcoming legislation, regulation, and public policy, not only in Canada but around the world. Impact and conformity assessment thus offer another example of an area worthy of exploration.

Criterion II: Promoting trustworthy AI critical to research or commercialization

Much of the discussions about AI governance have centred on values or principles that speak to public concerns about AI adoption, that is, the promotion of trustworthy AI. 111 Some widely shared concerns about AI have been addressed through the ethics declarations or principles. For example, the OECD AI Principles, Singapore's AI Verify, and the U.S.'s Blueprint for an AI Bill of Rights all list similar ideals. It may thus be helpful to use such applications as proxies to identify AI uses about which the public might have heightened reticence.

Criterion III: Recognized Canadian leadership

Identifying areas or initiatives in which Canada already commands leadership can reveal AI applications where Canada could drive the global conversation and for which Canada could ensure standards are aligned with its values and commercial interests.

Projects under Canada's Global Innovation Clusters provide useful guidance on this front because its goal is to leverage Canada's existing areas of strength—such as technology in British Columbia, advanced manufacturing in Ontario, and supply chains in Quebec—by applying emerging technologies to these sectors. ¹¹² Using AI to address climate change is another area in which Canada could demonstrate leadership. Considering its commitment to addressing climate change and established leadership in clean technology, Canada could make significant contributions as a proponent of the use of AI in the context of climate change.

This criterion focuses SCC'S AI and Data Governance Standardization Collaborative on areas in which Canada can leverage its existing assets, thus saving resources, capitalizing on existing expertise, and furthering its role as a world leader in various sectors, all while driving standards development at the international level.

31

¹¹¹ Neudert, Lisa-Maria, et al. "Global Attitudes Towards AI, Machine Learning & Automated Decision Making," Oxford Commission on AI & Good Governance, 7 Oct. 2020, online (oxcaigg.oii.ox.ac.uk).

[&]quot;Global Innovation Clusters," ised-isde.canada.ca.

Examples of the criteria in use

Table 3 (below) outlines three areas that fulfill the criteria, providing an example of the methodology in use.

Example	Identified in important legislation, regulation, or public policy	Promoting trustworthy AI critical to research or commercialization	Recognized Canadian leadership
Regulatory technology	The UK's Financial Conduct Authority and Competition and Markets Authority have started to develop and test Al-based regulatory technology, such as audit and assessment technologies	Informed by principles of human-centred values and fairness, transparency and explainability, and oversight accountability	Proposals for regulatory markets for AI; ¹¹³ work on leadership in regulatory technology ¹¹⁴
Credit scoring	Classified as high risk in the AI Act and the AAA	Engages principles of human- centred values and fairness, transparency, and explainability	Rapid growth of the fintech sector, especially in Toronto 115
Climate change	International treaties on climate change and national government positions on climate change	Engenders principles of human- centred values and fairness, security and safety, robustness, and accountability	Federal budget and Global Innovation Clusters on Digital Technology, Ocean, and Scale AI

Table 3: Examples of use cases, selected using the proposed criteria.

Additional key examples for the Collaborative's consideration

Other areas in which Canada can lead standards development include the financial sector (for example, automated lending), which is the focus of the AIMS pilot; healthcare (for example, automated diagnosis and treatment), a clear priority for the implementation of trustworthy AI; employment (for example, automated employment decision tools), in which the actions of other jurisdictions, such as New York City, can provide insight; and agriculture (for example, automated pest and soil-defect detention), an area in which Canada's Global Innovation Clusters have indicated interest.

Further, an overview of the international AI regulation and standardization landscape reveals many governments allocating particular focus to risk management, bias, and quality management. Promoting the development and adoption of these areas will be integral to maintaining cohesion with international frameworks and markets. Building out standards that focus on these concepts will help serve the global AI community. For example, the Collaborative may consider identifying priorities for businesses in areas where industry is struggling with risk management.

¹¹³ Clark, Jack, and Gillian K. Hadfield. "Regulatory Markets for Al Safety," arXiv, 11 Dec. 2019, online (arxiv.org).

^{114 &}quot;Canada's road to RegTech adoption," KPMG, online (home.kpmg); "Solving Regulatory Challenges Through Technology Innovation," Canadian Regulatory Technology Association, online (canadian regtech.ca).

^{115 &}quot;The Rise of Fintech in Canada: Home to a Vibrant Innovation Ecosystem," The Fintech Times, 30 Mar. 2022, online (fintechtimes.com).

CONCLUSION

Attempts to regulate AI have several competing characteristics, creating a noisy governance space in which it is difficult to discern a clear signal. The race for regulators to catch up with the rapid pace of innovation has created a fragmented and fraught landscape. Stakeholders around the world have generally agreed that standards and conformity assessment are vital tools for AI governance. Yet, standards alone are not a panacea for international AI regulation and cohesion. Further, to be effective, international standards will require additional work by stakeholders, if they are to be implementable in specific sectors or applications.

Still, given the importance stakeholders are placing on these standard and conformity assessments, it will be important for the Collaborative to understand and connect with standards-setting and conformity assessment-related initiatives. Doing so will allow Canada to make its voice heard on the world stage in regard to the development of standards, assuring conformance to them, and help drive effective regulation. There are several ways in which Canada can not only help determine the content of standards and conformity assessment, but also drive their development in efficient and effective ways.

In particular, the Collaborative should consider how SCC and other actors might undertake initiatives, such as the current AIMS pilot, that work to translate general international standards into concrete, industry-specific guidelines that organizations can understand and implement. A shift towards performance-based and/or product-centric standards and conformity assessment schemes, approaches, or programs in particular use cases will greatly increase their effectiveness as a governance tool.

By continuing to invest in mutually beneficial bilateral partnerships with major players, encouraging Canadian participation in standards development, and implementing innovative standards initiatives such as conformity assessment pilots and sandboxes, Canada can play a role in improving international cohesion, while advancing important domestic interests. Developing and supporting internationally accepted conformity assessment programs under the network of international accreditation bodies will reduce the risk to Canadian AI businesses and customers by ensuring that an accredited certificate or validation/verification statement may be relied upon and recognized anywhere in the world.¹¹⁶

It is vital that the Collaborative, in its work, understands and recognizes Indigenous sovereignty. Input and participation from Indigenous stakeholders cannot be a box-ticking exercise, and the Collaborative should undertake to meaningfully engage with the issue of data sovereignty by continuing work already started by the DGSC. ¹¹⁷ Further, the Collaborative should be mindful of the interests of SMEs, given that SMEs comprise the majority of all Canadian AI businesses despite their relatively small scale of resources to dedicate to standards development and implementation.

This paper provides an entry point for the Collaborative to better understand the international AI policy and standardization landscape, including the opportunities available to Canada. The recommendations addressed here are not an exhaustive list, but rather considerations of particular priority and relevance from which the Collaborative can begin to develop a strategy and a roadmap. As demonstrated, Canada is well situated to play a meaningful role in developing international standards, increasing understanding of such standards, creating conformity assessment tools and mechanisms to enable concrete implementation by Canadian organizations, and protecting and promoting societal interests.

¹¹⁶ International Accreditation Forum, https://iaf.nu/en/about/roles-and-benefits/

¹¹⁷ "Canadian Data Governance Standardization Roadmap," <u>scc.ca</u>.

APPENDIX I: METHODOLOGY

Desk research

To compile the necessary insights for this paper, research was conducted using qualitative methods. Firstly, a broad sweep of the international AI regulation and standardization landscape was undertaken, using both primary and secondary sources. A list of all potentially relevant regulatory initiatives, standardization initiatives, standardization bodies, and major initiatives related to AI was compiled for each of Canada's closest trading partners, as well as other main players in the field of AI. From this list, only those initiatives most directly relevant to the Collaborative's fulsome understanding of the international AI landscape were selected for more rigorous study and explanation. From this starting point, additional necessary areas of research were identified, and further research was conducted on all topic areas addressed in this paper, again using both primary and secondary sources.

Interviews

A list of key AI stakeholders both within and outside of Canada was generated, and interview requests were sent to individual stakeholders by email. Included amongst those interviewed were individuals with knowledge of or positions at AI start-ups, large technology corporations, the Government of Canada, intergovernmental organizations, national standards bodies, Indigenous research and advocacy organizations, and international standards bodies. Interviews were conducted via virtual video meeting platforms, with each interview lasting approximately one hour. Interviewees were asked between 8 and 12 questions on their area of expertise, with a question list available prior to each interview, upon request. Each interview involved the preparation of relevant questions, note-taking during interview sessions, and a team debrief to pinpoint the major insights to be noted, further researched, and included in the white paper. A total of 16 interviews were conducted (see Appendix II).

APPENDIX II: CONSULTATIONS

We are grateful to experts from the following organizations who generously offered their time for consultations as part of our research:

- Armilla Al
- British Standards Institute
- Canadian Institute for Advanced Research
- CIO Strategy Council
- Credo Al
- Enterprise Singapore
- European Committee for Standardization and European Committee for Electrotechnical Standardization
- First Nations Information Governance Centre
- Global Affairs Canada
- Google (former employee)
- Innovation, Science, and Economic Development Canada
- Institute of Electrical and Electronics Engineers
- ISO/IEC JTC 1/SC 42
- Manulife
- Microsoft
- National Institute of Advanced Industrial Science and Technology
- National Institute of Standards and Technology
- Organisation for Economic Co-operation and Development
- Responsible Artificial Intelligence Institute
- Standards Australia
- Standards Council of Canada