Digital Policy Hub — Working Paper

Standards as a Basis for the Global Governance of AI in Research

Matthew da Mota

About the Hub

The Digital Policy Hub at CIGI is a collaborative space for emerging scholars and innovative thinkers from the social, natural and applied sciences. It provides opportunities for undergraduate and graduate students and post-doctoral and visiting fellows to share and develop research on the rapid evolution and governance of transformative technologies. The Hub is founded on transdisciplinary approaches that seek to increase understanding of the socio-economic and technological impacts of digitalization and improve the quality and relevance of related research. Core research areas include data, economy and society; artificial intelligence; outer space; digitalization, security and democracy; and the environment and natural resources.

The Digital Policy Hub working papers are the product of research related to the Hub's identified themes prepared by participants during their fellowship.

Partners

Thank you to Mitacs for its partnership and support of Digital Policy Hub fellows through the Accelerate program. We would also like to acknowledge the many universities, governments and private sector partners for their involvement allowing CIGI to offer this holistic research environment.

About CIGI

The Centre for International Governance Innovation (CIGI) is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

<div style="border:1px solid green;">

Key Points

- Artificial intelligence (AI) risk in the research sector is an issue that goes beyond individual institutions or regions and will require international governance to ensure the preservation of an independent and reliable global research sector.

- While many types of governance and safety tools will be necessary, the use of international standardization would be a useful initial phase that could allow for rapid and inclusive governance on which other frameworks could build.

- Standardization allows for the kind of stakeholder engagement and consensus building necessary to establish effective international governance on AI among disparate stakeholders.

- The existing Canadian standard for AI/machine learning (AI/ML) implementation in research institution — CAN/DGSI 128, currently in development by the Digital Governance Standards Institute (DGSI) — could be proposed and taken under review to be adopted as an international standard.

- International standardization for AI implementation in research institutions could help to build global consensus on preserving and protecting research institutions and their information from AI risk and could help to support other efforts at global AI governance outside of the sector.

</div>

Introduction

This working paper will argue that standardization is the ideal route for establishing robust and adaptive AI governance for the research sector internationally. This is mainly because of the ability for standardization to engage multiple stakeholders with different interests while ensuring accountability and robust requirements through conformity assessment and certification within a standard.

The first section of the paper will begin with an evaluation of the primary global risks posed by AI use within research institutions. The second section will discuss how these global risks require global governance structures and what that might look like in the research context. The third section will address some of the primary concerns or critiques of standardization as a method of governance and potential paths to ensure those shortcomings will not affect efficacy and robustness of an international standard for AI use in research. The fourth section will review the existing landscape of international standards for AI, the process of international standard development through the International Organization for Standardization (ISO) and other key institutions and organizations that might play a role in an ISO standard for AI in research. The fifth section will detail the essential elements of any ISO standard for AI in research. The final section will provide conclusions and recommendations for institutions to develop the CAN/DGSI 128 standard into an international standard.[1]

---

1   See https://scc-ccn.ca/standards/notices-of-intent/digital-governance-standards-institute/machine-learning-and-ai-0.

The Global Risks of AI In Research Institutions

One of the greatest challenges of developing meaningful governance for AI use within research institutions is the diffuse nature of the sector. Every university, library, archive or research centre has its own policies and interests, including the many centres, libraries and departments within each university that can have different priorities and policies. Provincial and national governments also have different relationships to higher education and research, as well as different priorities for research and higher education.

In Ontario, for example, new legislation in the form of Bill 194, the Strengthening Cyber Security and Building Trust in the Public Sector Act,[2] provides cybersecurity and AI guidance for public institutions, including universities, while the proposed national Canadian legislation for AI, the Artificial Intelligence and Data Act (Government of Canada 2023), does not mention the governance of research institutions at all. Within one nation, the national and provincial priorities are vastly different. Compounding the complexity in this example are the different funding relationships between research institutions, the federal government and provincial/territorial governments, which provide significant funding to universities and research institutions but through different channels and for different purposes. This network of funds, interests, priorities and conflicting perspectives on governance make for a fragmented governance and policy landscape.

These challenges in coordinating governance and engaging stakeholders across contexts are compounded when considering the task of developing governance for AI in research institutions internationally. International cooperation introduces the added complexity of bridging different legal systems, values and ethical frameworks, as well as potential barriers in the form of international trade agreements or foreign policy concerns. Some might consider turning to existing international governance bodies such as the United Nations to develop international AI governance, but those institutions are limited in their capacity to compel states to follow governance frameworks and are experiencing decreases in trust from many nations for various reasons.[3] This issue is particularly acute when considering the global risks of AI and potential global governance paths to ensure collective safety and benefits of AI in an increasingly polarized world.

Despite these challenges, it is imperative that such a coordinated effort for international governance of AI use in research is established. As described in the first of a series of working papers by the author exploring this topic, modern research institutions are not siloed or cloistered but instead are parts of vast networks of information sharing (da Mota 2024b). These networks involve government-sponsored data and information repositories, institutional systems, private repositories owned by large copyright-owning companies, physical interlibrary loan services and hybrid systems through which information and data flow. The interdependent nature of research networks makes AI risks in research institutions

---

2   Bill 194, *An Act to enact the Enhancing Digital Security and Trust Act, 2024 and to make amendments to the Freedom of Information and Protection of Privacy Act respecting privacy protection measures,* 1st Sess, 43rd Leg, 2024 (first reading 13 May 2024), online: <www.ola.org/sites/default/files/node-files/bill/document/pdf/2024/2024-05/b194_e.pdf>.

3   A recent poll found that the majority of Canadians do not trust the United Nations (Chang 2024), while UN Secretary-General António Guterres has acknowledged directly that such institutions are eroding in terms of efficacy and public trust and are in need of reform (UN Secretary-General 2024).

a global issue. The main risk vectors of AI in the sector are information loss, poisoning or degradation; intentional or accidental dis-/misinformation within information repositories; AI monopolies in the sector exerting significant influence on content and data access; data scraping to support training for research-specific large language models (LLMs);[4] and overreliance on tools leading to a potential loss of diversity and quality in research.[5] These risks are already concerns on a small scale that can be significantly amplified if they spread beyond a single institution, nation or region. Further, the understandable desire to preserve academic freedom often means that governments are hesitant to impose restrictions or regulations on research institutions and that institutions are often reluctant to request such guidance. The research sector is poised to be a high-risk but low-governance space, the worst combination of circumstances in which AI harms could proliferate globally.

Global Risks Require Global Governance

Experts on AI do not agree on when artificial general intelligence (AGI)[6] and artificial superintelligence (ASI)[7] will be achieved or if it will happen at all.[8] However, given that many acknowledge the potential significant risks posed by AGI or ASI, it behooves humanity to collaborate on governing AI in such a way that alleviates these potential risks and finds ways to harness AI for the public good. Duncan Cass-Beggs et al.'s (2024) recent discussion paper for the Centre for International Governance Innovation's Global

---

4   Enhanced AI search tools with built-in LLM suggestions and research support tools, such as Perplexity AI, have come under criticism for data and content scraping from news companies such as Condé Nast (Hindy 2024). This is a common practice among LLMs and particularly important for research tools such as Perplexity that need to be able to reference and index information, as well as to support their generative models' training. With research-oriented tools where scraping may occur on a university library website, the content might be owned by a copyright owner or shared under a specific content-sharing agreement, further complicating the issue. The complex interrelations between publishers, scholars, copyright owners and institutions add additional layers of difficulty in discerning plagiarism or data theft and in determining the best path for effective governance.

5   See references to scientific papers that were clearly created using generative AI, which has led to inaccurate and completely fabricated elements of some research papers (Park 2024). The role of generative AI in research has led to extensive discussion within the scientific community (Elali and Rachid 2023); many journals have even rewritten their submission guidelines to more tightly control the use of AI in writing or supporting research that is published in their journals (Harker 2023).

6   The definition of AGI is variable but can generally be defined as an AI system that can serve general purposes across multiple tasks (not only an LLM or image generator but multiple areas of focus) and can match or exceed humans' capacity in many of these tasks. Some argue that this intelligence must be embodied in a physical form capable of interacting with the world for AGI to be achieved (Gopalakrishnan 2022), while others disagree on this point (Tan and Jaiswal 2023). For a history and breakdown of the term and conception of AGI, see Lauren Leffer's (2024) primer in *Scientific American.*

7   ASI is a somewhat newer term than AGI and refers to the intelligences that might come after AGI is achieved, when scientists can place multiple AGI researchers on the task of improving algorithms and other elements of AI to create intelligences far exceeding the abilities of humans. There are several perspectives on the plausibility and route that ASI might take, but this working paper's use of the term is heavily informed by Leopold Aschenbrenner's (2024) essay series, Situational Awareness.

8   See Aschenbrenner's (2024) essay series turned e-book *Situational Awareness* (2024), as well as his interview on the *Dwarkesh* podcast (Patel 2024) for one perspective on how AGI and ASI might soon be achieved and how we might govern the risks involved. Scholars such as computer scientist Gary Marcus have also voiced concerns about the existential risks of AI, while arguing that the current LLM focus is not the path to AGI (see his Substack for articles on related topics: https://garymarcus.substack.com/). Computer science professor and Chief AI Scientist for Facebook AI Research Yann LeCun has spoken out extensively in response to critics such as Marcus about championing the pursuit of achieving AGI for the benefit of humanity in spite of the potential for existential risk; see Perrigo (2024) as an introduction to some of these ideas..

AI Risks Initiative explores the potential paths for such governance, highlighting the importance of states' broad representation even if there are disagreements about values and interests.[9] In short, we might interpret this to mean that even if the United States and China disagree about human rights, democratic values, sovereignty of other states and cultural matters, these world leaders in AI development must nevertheless develop a working agreement that can meaningfully govern AI and mitigate risks. Although some might argue against collaboration with nations that are not perceived to be suitably democratic or share Western values, there seems to be a growing consensus that some form of global collaboration must be achieved to ensure AI safety.

Some venture capitalists and big tech investors have argued that any kind of regulation will stifle creativity and the development of AI and that we should instead allow the market and innovation to govern the path of AI.[10] This perspective, often framed within discussions of "techno-optimism," appears to be more of a kind of techno-messianism in which blind trust is placed in abstract conceptions of "the market" and "science" that are not rooted in fact and do not reflect the geopolitical reality of AI that we see unfolding. It seems that a careful appraisal of the expert opinions and developing dynamics in global markets suggests that some form of rapid governance, however difficult, is necessary.

In the context of research institutions, global governance is both harder, because there are far more individual interests at multiple levels, and easier, due to histories of collaboration and shared values of pursuing and furthering research that transcend national interests. Though in Canada we have seen much discussion about Chinese interference in research and politics,[11] Canadian research institutions still share strong relationships with Chinese institutions, and meaningful and fair collaboration persists.[12] Research institutions have a base for collaboration that could serve as a pilot for more comprehensive global governance.

The significant risks AI poses to the research sector mean that any effective governance framework for AI, including those proposed earlier in this series, must be interinstitutional, and eventually international, to ensure data and information security and quality throughout vast research networks. Moreover, any governance must have ethical and security-focused underpinnings while also remaining inclusive and avoiding partisan and protectionist tendencies. On one hand, academia and research have long

---

9   Cass-Beggs et al. (2024) also explore several potential paths for global AI risk governance as a primer to initiate global cooperation on the topic. The position of the initiative seems to be that regardless of disputes about whether these risks are likely or unlikely, or plausible or implausible, any level of potential risk warrants serious cooperation to establish governance. The paper also focuses on the importance of global coordination to realize the benefits of AI globally, which is another important perspective that goes hand in hand with the risk discussion.

10   Perhaps the most prominent of these arguments, which generated extensive debate on social media platforms, is "The Techno-Optimist Manifesto" by investor Marc Andreesen (2023).

11   Canadian news has been filled with discussions of potential Chinese interference in politics (Paas-Lang 2024), as well as discussions of Chinese espionage in research contexts, namely the alleged sharing of information by two Chinese nationals working at Canada's infectious diseases lab in Winnipeg (Tunney 2024). In response to these issues and other increasing fears of Chinese espionage and interference, the Canadian government has released several guidance documents on research security that seem to focus primarily on Chinese interference (see Canada's research security information landing page: www.canada.ca/en/services/defence/researchsecurity.html), while also clearly aiming to maintain the benefits of Canada's extensive research connections to China.

12   The existing rules to combat foreign interference in research are very specific and limited to a list of Chinese and/or mostly military and espionage-related institutions with which researchers have been instructed to avoid collaboration. Despite this list of institutions, however, there is no permanent ban placed on collaboration but rather a kind of cool-down or grace period expected between time spent at one of the institutions on the list and assuming a position in a Canadian research institution.

been spaces for collaboration across borders, and it is imperative that the sector does not become a battleground for rising international tensions. On the other hand, it cannot be ignored that research has become politicized and is now a space for intellectual property theft, espionage, and influence and manipulation by state and non-state actors. It is essential that any international agreement also respects and protects sovereignty and research independence and does not enable or embolden foreign interference of any kind.

With the proposed national standard for AI in research, CAN/DGSI 128, in the drafting stage, there is an opportunity to bridge the gap between conflicting Canadian interests to establish a nationally standardized framework for AI use in research institutions.[13] The DGSI and its parent organization, the Digital Governance Council (DGC), are recognized as official organizations that can submit a Canadian national standard to international bodies such as the ISO to be reviewed and adopted as international standards. Rather than the entire process having to be started from scratch at the international level, the CAN/DGSI 128 standard could be adopted as is by the ISO or some other international organization, pending a successful review.

The EU AI Office's efforts to develop the EU AI Act into a general purpose AI code of practice is the first attempt to turn AI legislation into a practical guide for institutions and organizations. This process will involve the review and possible adoption of existing standards and codes of conduct, including the Group of Seven (G7) Hiroshima AI Process Code of Conduct, along with some of the language and tools from the Organisation for Economic Co-operation and Development (OECD) (Allen and Adamson 2024). Although the EU AI Act does not mention research institutions, its code of practice will be an important guide for understanding how standards and other codes are braided together into practical guides for AI, especially at the international and regional levels. If agreements such as the EU AI Act rely partly on existing standards, then it is essential that the research sector put a standard into place to provide guidance on this subject.

More comprehensive and government-driven solutions will be necessary, especially ones that address the significant role of AI companies in developing AI and the potential for significant risks from AI beyond those involving information or research. However, international governance standards generally, and in the research context specifically, can serve as a base for further governance development. It is also important to note that standards governance does not constitute regulation but is instead an opt-in regime that not only supports safety and mitigates risk, but also promotes the successful use of whatever is being governed (in this case, AI). In the research sector, such promotion of innovation and development could help to protect data and the core values of research institutions from AI risks, while also establishing a plan for mobilizing and making use of significant, previously unused or undiscovered data within institutions that could support further AI research.

---

13 See the second paper of the author's working papers on this topic, "University of Toronto Libraries: A Case Study for AI Governance in Research Institutions" (da Mota 2024c), for more in-depth analysis of the priorities for the CAN/DGSI 128 standard. It is still true that not one of the proposed or established pieces of AI legislation (including the recent EU AI Act) mentions research institutions' use of AI as a priority or area of risk, despite the sector being at the centre of AI development (ibid.). It is also still true that research institutions are central to the preservation and furthering of human knowledge, history and culture, and must be protected from both potentially malfunctioning AI systems and predatory practices that might use data-heavy AI systems such as LLMs as a means of accessing and controlling proprietary institutional data (ibid.). These realities highlight the importance of this standardization both nationally and internationally.

Critiques of Standardization

Standards are often framed by standards institutions as apolitical and independent, which is often true. However, there are several potential criticisms of standardization that might derail the creation of an AI standard for the research sector that should be addressed to maximize the legitimacy and transparency of the process.

A valid criticism of standards development is that it is not as apolitical or objective a process as some might believe (Solow-Niederman 2023). This argument is based on the fact that the initiators and proponents of a standard are often affected by its outcome; their corporate affiliations or other interests might somehow be improved by passing the standard in question. While it is true that the potential for manipulation or bias exists in the standardization process, the general practices of stakeholder engagement in the drafting and review process work to balance any biases that might exist in the resulting standard.

When it comes to international governance of AI in research, some developing countries or non-Western powers, such as China, might fear that a Western-driven standard could act as a Trojan Horse, carrying American and/or Western values into a global agreement. In the same vein, the United States, the European Union and other Western powers might be suspicious of China's or Russia's involvement in such a process. In both cases, the fear is that the standard will become a means of forcing political, economic, or social interests upon other nations, turning negotiations into a proxy battleground for discussions of values and political perspectives. While this is a potential risk, it is also a necessary one, given the high stakes discussed earlier in this paper. A focus on consensus among members of the review committee would help to alleviate these concerns, as would the requirement that the standard undergo regular review, giving states a chance to evaluate and remedy any built-in biases it over time. In addition, if the committee were comprised of independently selected practitioners from research institutions, this would ensure a focus on research priorities and not political interests.

On one hand, the voluntary nature of standards means that many different interests across international research contexts can be represented in the development of governance, engaging many stakeholders. On the other hand, voluntary codes are far from perfect and can fail to hold institutions accountable, serving instead as a kind of ethics washing for AI (da Mota 2024a). The value of an international standard for AI in research comes only with the assurance that the standard will be effective. Any standard developed through the ISO would need to include a conformity assessment program so that the benefits of voluntary agreements, namely, openness and buy-in from institutions, are not made toothless or ineffective by a lack of accountability.

An international standardization process could use the ISO's "Conformity assessment — General principles and requirements for validation and verification bodies" (ISO 2019) to build a comprehensive governance framework including a governance standard, establishing a conformity assessment regime

and developing a process for certifying professionals as competent to assess an institution's conformity to the standard. Under ISO 17029 (2019), a nation (Canada), national association (DGC or DGSI) or industry association (an international research organization) could oversee and own the program that would include all three of the standardization elements. This allows for any international institution with the capacity and mandate to do so to own the whole program of standardization, including the conformity assessments and reviews. In addition to using the basic standard initiated as CAN/DGSI 128, the body that takes on oversight of the program can pursue the development of further documentation and standards — facilitated by an international standards organization such as the ISO or the Institute of Electrical and Electronics Engineers (IEEE) — to help govern emerging issues such as specific technical requirements, tool-specific programs or additional certification and testing requirements.

Any conformity assessment could be done through a peer-reviewed process in which certified individuals from member institutions of the standardization program would conduct an evaluation of another member institution based on the established conformity assessment plan. This peer-review process would be in keeping with the collaborative and peer-reviewed work already occurring within academic research and would help to focus the standardization of the sector on a collective effort by institutions to self-govern fairly and democratically, and to make use of existing relationships and resources rather than hiring third parties outside of the sector to conduct assessments and reviews. It would also help to build confidence that the assessments were independent by involving multiple reviewers from other institutions in each assessment, while also maintaining the confidence that assessors are professionals from the same field and thus able to understand and engage with the unique dynamics and concerns of academic and research environments.

Standards are not a perfect governance tool, but they are one of several tools that must be utilized to develop an effective and representative governance platform that can represent and hold accountable thousands of institutions around the world. This tool would be a baseline on which other standards and agreements could be developed, as well as a direct and useful way to begin this essential standardization process and support similar efforts in future.

The International Standards Landscape for AI in Research

The ISO is the largest and oldest international standards body and has a strong international reputation for building effective standards that meet the needs of the market and the stakeholders engaged by a standard. The ISO's standards development process is comparable to most other organizations and involves the following stages: creating a proposal; assembling an expert drafting team to draft the standard; consulting broader committees and implicated organizations or experts on the draft; reviewing and revising it; achieving consensus and final approval; disseminating the standard and having organizations sign on;

and further ongoing review and revision.[14] When a standard contains auditable requirements, other building blocks covered under the standardization system need to be designed and implemented. A conformity assessment program or scheme should be put in place whereby organizations' adherence to the standard is evaluated and recommendations for improvement or revocation of membership for failure to conform to the standard can be delivered. Standardization can also include a certification process that ensures proper training and standards for individuals using the technology governed by the standard, systems that deploy this technology and organizational bodies that use it.

The existing ISO standards, "Information technology — Artificial intelligence — Guidance on risk management" (2023a), "Information technology — Artificial intelligence — Management system" (2023b) and "Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)" (2022), are examples of ISO standards for AI that focus on different aspects of the technology. They are important seed documents for future AI standards for research and establish a precedent for the kind of broad global standards for AI that can focus on specific risk areas or functions of the technology. The ISO's role complements the broader community-driven grassroots project by offering an independent and neutral framework for constructing and implementing the standard itself. It is also worth noting that the ISO is on the American Library Association (ALA)'s list of international information organizations, recognizing the ISO's role in supporting and furthering information policy and governance alongside other organizations more specifically focused on research institutions and universities.[15]

Given the broad remit and importance of this proposed standard and the implications of governing research specifically, other standards institutions, including national organizations (the National Institute of Standards and Technology and the DGC), regional organizations (the European Committee for Electrotechnical Standardization in the European Union) and international industry or profession-specific organizations (the IEEE), might also be necessary to help promote and support the development of the standard. With the Canadian standard, as well as strong engagement and uptake from institutions, there will be a precedent and a path for the DGC and the DGSI to lead this international standard through the ISO or the IEEE, where it would be submitted as a national standard, reviewed as an international standard and then either edited based on the review or accepted as is. The focus of the DGC and DGSI standard currently is not only on the Canadian context but also on forming a standard that is fit for purposes across different kinds of institutions globally; while there is a tight focus on research-specific issues related to AI, it comes with a broad scope that is suitable for non-Canadian institutions as well. It would be useful and important for international organizations, or at least their Canadian partners to be involved in the national standardization process now, so that the Canadian standard is suitable to be adopted by an international standardization body. As Canadian research institutions and universities involved in the process are very interconnected and internationally focused institutions, the national standard will be focused on the most important

---

14 See the ISO guidance on developing standards for more on all aspects of the standardization process through their program: www.iso.org/developing-standards.html. Also see the IEEE developing standards page, which outlines a similar process: https://standards.ieee.org/develop/.

15 See the ALA international page: www.ala.org/aboutala/offices/iro/intlassocorgconf/international.

issues for a modern, world-class and globally networked research institution and the many challenges it will face with adopting AI.

Outside of standards organizations, there are multiple international organizations that represent the interests of research institutions and could help to assemble drafting teams, engage stakeholders, elicit feedback, conduct conformity and efficacy assessments, and implement training or certification processes for their member organizations. The key organizations representing research institutions internationally are the International Federation of Library Associations and Institutions, the International Alliance of Research Library Associations, the International Association of University Libraries and the International Council on Archives. These organizations represent broad coalitions and consortia of research institutions, archives, libraries and universities to varying degrees and could potentially have the mandate to speak for their membership on the development of an AI governance standard, or could be entrusted through a members' vote to do so if necessary. Other important organizations that develop standards or have a focus on the preservation of knowledge and promotion of education and research are the International Labour Organization,[16] which focuses on global labour rights and research; the United Nations Educational, Scientific and Cultural Organization (UNESCO), which focuses on protecting human heritage, including extensive work to protect and preserve important information resources and data; and the World Intellectual Property Organization, which has influence over questions of copyright and the protection of information in the research context.

These institutions would play the role of representing the interests of their members, engaging with their communities and others impacted by the standard; they could also administer and promote the completed standard to ensure continuing representation and effective assessment of efficacy, moving forward in an advocacy role to balance the ISO's neutral organizational and administrative role. It would be necessary to evaluate the different kinds of organizations to assess whether they would play an advocacy or advisory role, such as the United Nations or UNESCO, or whether they have the mandate to speak on behalf of their members, as might be the case for some regional, national or international consortia. These institutions will be essential in ensuring the viability of the standard as one that institutions will trust and be willing to adopt, as well as in evaluating its effectiveness and conducting broad reviews to update or redraft the standard if necessary to respond to changing needs.

Institutions such as the G7 (2023), the UN High-level Advisory Body on Artificial Intelligence (2024), UNESCO (Miao, UNESCO and Holmes 2023) and the OECD (2023) have guidance on AI that might be useful to the standard development process, and their international governance bona fides might lend themselves to supporting the process. But these institutions also come with their own baggage and are not research focused, so it will be a balancing act for any of them to maintain their role in supporting and promoting the process without leading or controlling it.

Beyond the many institutions developing and consulting on the ISO standard for AI in research and those supporting this process, this standardization effort would also require strong pushes from governments and institutions to support the propagation, adoption and adherence to the proposed international standard, thereby ensuring that many institutions

---

16 Ibid. Listed by the ALA as an important organization because of their work on labour rights standardization and the preservation and dissemination of related work.

sign on to and accept the authority of the standard. Canada's initial role in developing the first standard for AI in research institutions,[17] as well as the nation's likely role in pushing the standard toward international development and adoption through the ISO, will rest on the country's strong position in research, the influence of its standards organizations such as the DGSI and the DGC, and its international reputation as a peacebuilder. However, it is imperative that Canadian institutions not be seen as sliding toward certain biased perspectives on AI and research, in order to remain neutral leaders on this project.

Essential Elements of the Standard

Drawing from the author's previous working papers on this topic and other important documents referenced throughout this working paper, there are several key concerns that must be addressed in any standard for AI implementation in research institutions.

- Any successful standard must engage as many stakeholders as possible and represent a diverse array of institutions globally. The standard must also have the support of international bodies before, during and after development to ensure its uptake and success among institutions.

- There are two main focuses of the CAN/DGSI 128 standard that might also shape an ISO standard for research:

  - how AI is used within the internal systems of an institution to collect, preserve, manage and distribute information; and

  - how AI is used in actual research, in both the development of AI tools for research and the adoption of existing tools to support research.

- There must be language on data protection, privacy and acknowledgement of open-science, open-source and open-access questions, as well as of copyright and cybersecurity as subtopics that are of essential importance to understanding AI governance.

- There must be consideration of the risks of monopolies on the sector and a consensus on independent research as a global good and how to maintain the autonomy of research institutions individually.

- Questions of foreign interference and security should be mentioned to the degree that research institutions might be implicated in those topics, though without straying from the research-specific mandate of the standard.

- A conformity assessment element must be included, referencing ISO 17029, including third-party assessments that should be done through a peer-review assessment program to keep costs down and to build trust and buy-in to the review system between institutions.

---

17   See https://scc-ccn.ca/standards/notices-of-intent/digital-governance-standards-institute/machine-learning-and-ai-0.

- This standard must be sensitive to and informed by existing international rights and laws to ensure that it supports compliance with the law, while also identifying gaps in existing frameworks to ensure a higher standard of protection.

- It must be acknowledged and maintained throughout the process that the purpose of the standard is to ensure the preservation of an independent and reliable global research sector.

There are other key ideas that will no doubt emerge in the development of the standard, but these issues are central concerns that should be built into the process from the beginning as they are essential to the basic functions of research institutions and cover the significant areas that shape AI use in the sector. The proposed framework for AI governance described in da Mota (2024b) should also should also be considered as a guiding set of principles for these standardization efforts.

Conclusions

This paper has proposed that an international standard to govern AI use in research institutions is an essential baseline to ensure the security and longevity of research globally, and also has the potential to support other more comprehensive global AI governance efforts in the future. The Canadian standard for its research institutions currently under development, CAN/DGSI 128,[18] could be adopted by an international standardization organization upon review of its suitability in an international context.

AI use and experimentation is likely to push research institutions into a new era of exploration and productivity that will hopefully benefit all of humanity. However, this technology also poses significant risks to research and to the essential human knowledge and heritage held and protected within research institutions in their many forms. Developing an international standard, that is agreed upon by experts in the field internationally could afford an essential protection to ensure that AI use in research only pushes knowledge forward and does not cause potentially disastrous outcomes. Beyond protecting knowledge and supporting new research, this kind of standard could serve as an example of the type of international cooperation and collaboration that can be undertaken and achieved on AI as well as other important topics.

Recommendations

- The Canadian federal government and provincial/territorial governments should support the adoption of the CAN/DGSI 128 AI standard to build a strong culture of compliance and collaboration around it.

- The Standards Council of Canada, along with the DGSI, the DGC and other Canadian standards organizations, should support and work to propose and pass the CAN/DGSI 128 as an international standard through the ISO or the IEEE.

---

18   Ibid.

- The current development of CAN/DGSI 128 must focus on international implications and challenges to ensure that it will be fit for purpose at the international level.

- Any future AI treaties or governance should take any ISO AI for research standard as a seed document and key thread to support and underpin other global AI governance plans.

The Canadian government (perhaps led by Innovation, Science and Economic Development Canada, the Social Sciences and Humanities Research Council, the Natural Sciences and Engineering Research Council or a combination of these bodies) and international organizations such as the Group of Twenty, the United Nations and the European Union must work to support and develop the necessary digital public infrastructure to support the success of the standardization process and the safe and beneficial development of AI tools. This includes digital infrastructure for data exchange and sharing; digital codes to govern other aspects of digital research that support and affect AI; and physical infrastructure and policy to support the significant energy needs of the sector in ways that are in keeping with climate goals.

Acknowledgements

About the Author

Matthew da Mota is a senior research associate and program manager for the Global AI Risks Initiative at CIGI, working to develop governance models to address the most significant global risks posed by AI and to realize the potential global benefits of AI in an equitable and sustainable way. He is particularly interested in how emerging technologies such as AI might shape information systems, ways of thinking and institutions over time and what kinds of collaborative research and governance models can address these challenges.

During his post-doctoral fellowship at CIGI's Digital Policy Hub, Matthew focused on AI policy and implementation in research and higher education, spanning institutional policies, governance standards, content and tool licences, and data governance frameworks. He continues to lead governance in the sector as chair of the expert drafting committee for a national governance standard for machine learning and AI implementation in research institutions with the DGSI.

Matthew holds a Ph.D. from the University of Toronto Centre for Comparative Literature.

Acronyms and Abbreviations

AGI	artificial general intelligence
AI	artificial intelligence
ALA	American Library Association
ASI	artificial superintelligence
DGC	Digital Governance Council
DGSI	Digital Governance Standards Institute
G7	Group of Seven
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
LLMs	large language models
ML	machine learning
OECD	Organisation for Economic Co-operation and Development
UNESCO	United Nations Educational, Scientific and Cultural Organization

Works Cited

Allen, Gregory C. and Georgia Adamson. 2024. "Advancing the Hiroshima AI Process Code of Conduct under the 2024 Italian G7 Presidency: Timeline and Recommendations." Center for Strategic and International Studies. March. www.jstor.org/stable/resrep58540.

Andreesen, Marc. 2023. "The Techno-Optimist Manifesto." Andreesen Horowitz, October 16. https://a16z.com/the-techno-optimist-manifesto/.

Aschenbrenner, Leopold. 2024. "SITUATIONAL AWARENESS: The Decade Ahead." Essay series. *For Our Posterity* (blog), June 14. https://situational-awareness.ai/.

Cass-Beggs, Duncan, Stephen Clare, Dawn Dimowo and Zaheed Kara. 2024. *Framework Convention on Global AI Challenges*. CIGI Global AI Risks Initiative Discussion Paper, June 25. www.cigionline.org/publications/framework-convention-on-global-ai-challenges/.

Chang, Joshua. 2024. "Most Canadians and Americans don't trust the UN, poll finds." *National Post*, April 20. https://nationalpost.com/news/canada/canadians-americans-dont-trust-united-nations-poll.

Da Mota, Matthew. 2024a. "Voluntary Codes of Practice for AI Lack Accountability." Opinion, Centre for International Governance Innovation, February 12. www.cigionline.org/articles/voluntary-codes-of-practice-for-ai-lack-accountability/.

———. 2024b. "Toward an AI Policy Framework for Research Institutions." Digital Policy Hub Working Paper. February 15. www.cigionline.org/publications/toward-an-ai-policy-framework-for-research-institutions/.

— — —. 2024c. "University of Toronto Libraries: A Case Study for AI Governance." Digital Policy Hub Working Paper. August 19. www.cigionline.org/publications/university-of-toronto-libraries-a-case-study-for-ai-governance/.

Elali, Faisal R. and Leena N. Rachid. 2023. "AI-generated research paper fabrication and plagiarism in the scientific community." *Patterns* 4 (3): 100706. www.sciencedirect.com/science/article/pii/S2666389923000430.

G7. 2023. "Hiroshima Process International Code of Conduct for Advanced AI Systems." October 30. https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-code-conduct-advanced-ai-systems.

Gopalakrishnan, P. G. Keerthana. 2022. "Embodiment is Indispensable for AGI." LessWrong, June 7. www.lesswrong.com/posts/vBBxKBWn4zRXwivxC/embodiment-is-indispensable-for-agi.

Government of Canada. 2023. "The Artificial Intelligence and Data Act (AIDA) – Companion document." Innovation, Science and Economic Development Canada. https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document.

Harker, Jennifer. 2023. "Science journals set new authorship guidelines for AI-generated text." Environmental Factor, March. Research Triangle Park, NC: National Institute of Environmental Health Sciences. https://factor.niehs.nih.gov/2023/3/feature/2-artificial-intelligence-ethics.

Hindy, Joe. 2024. "Condé Nast Sends Cease-and-Desist to Perplexity AI Over Data Scraping." *PCMag*, July 23. www.pcmag.com/news/conde-nast-sends-cease-and-desist-to-perplexity-ai-over-data-scraping.

ISO. 2019. "Conformity assessment — General principles and requirements for validation and verification bodies." ISO/IEC 17029:2019. www.iso.org/standard/29352.html.

— — —. 2022. "Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)." ISO/IEC 23053:2022. www.iso.org/standard/74438.html#lifecycle.

— — —. 2023a. "Information technology — Artificial intelligence — Guidance on risk management." ISO/IEC 23894:2024. www.iso.org/standard/77304.html.

— — —. 2023b. "Information technology — Artificial intelligence — Management system." ISO/IEC 42001:2023. www.iso.org/standard/81230.html.

Leffer, Lauren. 2024. "In the Race to Artificial General Intelligence, Where's the Finish Line?" *Scientific American*, June 25. www.scientificamerican.com/article/what-does-artificial-general-intelligence-actually-mean/.

Miao, Fengchun, UNESCO and Wayne Holmes. 2023. *Guidance for generative AI in education and research*. Paris, France: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693.

OECD. 2023. "Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI." OECD Digital Economy Paper No. 349. Paris, France: OECD. https://doi.org/10.1787/2448f04b-en.

Paas-Lang, Christian. 2024. "Foreign interference inquiry will be 'deeply uncomfortable' for Beijing, Canada's ambassador to China says." CBC News, January 20. www.cbc.ca/news/politics/foreign-interference-inquiry-china-canada-1.7088853.

Park, Alice. 2024. "AI Writes Scientific Papers That Sound Great — but Aren't Accurate." *Time*, February 20. https://time.com/6695917/chatgpt-ai-scientific-study/.

Patel, Dwarkesh, host. 2024. "Leopold Aschenbrenner – China/ US Super Intelligence Race, 2027 AGI & The Return of History." June 4, in *Dwarkesh,* podcast, 4:31:17. www.dwarkeshpatel.com/p/leopold-aschenbrenner.

Perrigo, Billy. 2024. "Meta's AI Chief Yann LeCun on AGI, Open-Source, and AI Risk." *Time,* February 13. https://time.com/6694432/yann-lecun-meta-ai-interview/.

Solow-Niederman, Alicia. 2023. "Can AI Standards Have Politics?" *UCLA Law Review* 71 (2): 230–45. www.uclalawreview.org/can-ai-standards-have-politics/.

Tan, Cheston and Shantanu Jaiswal. 2023. "The Path to AGI Goes through Embodiment." In *Proceedings of the Inaugural 2023 Summer Symposium Series 2023*, Singapore EXPO, Singapore, July 17–19, edited by Harold Soh, Christopher Geib and Ron Petrick, 104–8. Washington, DC: Association for the Advancement of Artificial Intelligence. https://doi.org/10.1609/aaaiss.v1i1.27485.

Tunney, Catharine. 2024. "Scientist fired from Winnipeg disease lab intentionally worked to benefit China: CSIS report." CBC News, February 28. www.cbc.ca/news/politics/winnipeg-lab-firing-documents-released-china-1.7128865.

UN High-level Advisory Body on Artificial Intelligence. 2024. *Governing AI for Humanity: Final Report*. September. www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.

UN Secretary-General. 2024. "At Davos Forum, Secretary-General Warns of Global Norms Collapsing, Highlights Need to Rebuild Trust, Reform Governance." Press release, January 17. https://press.un.org/en/2024/sgsm22109.doc.htm.