

Digital Policy Hub – Working Paper

# Toward an AI Policy Framework for Research Institutions

**Matthew da Mota**

Fall 2023 cohort

## 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.

Copyright © 2024 by Matthew da Mota

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

Centre for International Governance Innovation and CIGI are registered trademarks

67 Erb Street West  
Waterloo, ON, Canada N2L 6C2  
[www.cigionline.org](http://www.cigionline.org)

## Key Points

- Research institutions, for example, research libraries, archives and universities, are an essential part of society that are at high risk for artificial intelligence (AI) adoption within basic systems such as search and cataloguing functions.
- No existing legislation or policy framework addresses the research sector or research institutions as a priority.
- Given its basic societal function, it is imperative that comprehensive AI policy be developed for the research sector in a collaborative, multi-stakeholder process.
- Existing policies for AI in universities, ethical frameworks for library and archive associations, and existing international and national policies are relevant to help construct a policy framework based on existing norms, practices and values.
- The core values of preserving information, preserving context, promoting accessibility, ensuring sustainability and fostering transparency are the foundational principles on which a policy for AI use in research institutions should be built.
- The Canadian landscape is well placed to lead international cooperation on this issue.

# Introduction

Research institutions and the networks they form are essential foundations of society's information infrastructure. These institutions' role in collecting, preserving and disseminating information across sectors is crucial to protect cultural heritage and history, promote innovation, uphold the principles of open science and support the right to education. In addition to the qualitative impact of research institutions on society, they also represent a significant economic engine. U15 Canada (U15) — an association of the 15 largest research universities in Canada — states that their member institutions conduct roughly \$8.5 billion in research annually, hold 83 percent of all contracted private-sector research in Canada and hold 81 percent of Canada's university patents.<sup>1</sup> The U15 members and other research institutions in Canada also hold vast amounts of cultural, scientific and historical documentation that constitutes a vast portion of human knowledge and historical records. These institutions ensure a robust system of research and innovation for social good and provide a factual base on which civil society and governance can operate. Considering these factors, it is evident that universities and other research institutions across Canadian society represent an unmatched body of knowledge, playing a crucial role within society.

Recent international discussions of AI governance have identified research in general (alongside education) as an important area for AI governance, but this remains an

---

<sup>1</sup> See <https://u15.ca/>; see also [www.univcan.ca/](http://www.univcan.ca/) for more statistics on the research impact of Canadian universities.

ill-defined risk area that has not been included in any significant legislation.<sup>2</sup> Any substantive discussion of the research sector and AI has focused only on the educational function of the institutions or the role institutions play in driving new AI research,<sup>3</sup> not considering the ways AI might affect how research institutions function or the potential risks of integrating AI into the basic systems of these institutions. Research institutions are an essential part of education, but they play the parallel (and at times more important) role of storing and making accessible, vast amounts of knowledge on which most academic and industry research is based, at least in part. The exclusion of research institutions from substantive AI legislation is, to some degree, a result of governments not wishing to interfere with academic freedom and research innovation. However, despite the need for academic freedom, these institutions play an essential role in society and any AI use in the sector must be guided by comprehensive policy. It may be the case that government legislation is not appropriate for various reasons. If so, it is necessary for the institutions in question to develop sector-wide policies to protect them from AI risks and ensure a coordinated and effective plan for yielding the maximum benefits of AI use in research institutions with governmental support.

Although not explicitly recognized in most discussions of AI governance, research institutions are a high-risk sector for AI use. The European Commission's definition of high-risk AI highlights several sectors including "critical infrastructures," "educational or vocational training," and "essential private and public services" (European Commission 2023), which research institutions fall under. Research networks constitute part of the critical infrastructure supporting research and housing vast amounts of data and information. Most of the research institutions mentioned are directly involved in supporting education and access to education. Finally, access to reliable information is an essential public service that research institutions deliver.

The imperative to develop policy for AI use in research institutions is due to the essential function of research institutions and research networks in society. As each of these institutions has a different mandate and serves different publics, this framework must address the most fundamental values and goals of all research institutions. Government should play a role in regulation, but in the interest of expediency and utilizing the expertise in the sector, policy development and implementation should be a collaborative process between existing research institutions and the associations that connect them, innovating by using existing ideas rather than reinventing the wheel.

This paper will survey existing policies relevant to AI use in research institutions. It will then explore the rewards and risks of AI use in research institutions and, considering these risks, will identify other relevant frameworks and documents that should inform policy development alongside formal legislation and policy. The paper will then propose a foundational framework for AI use across research institutions and next steps for policy makers.

---

2 See United Nations Educational, Scientific and Cultural Organization (UNESCO) (2023), which highlights research but does not specifically refer to how research institutions will use AI within their critical systems and on a more day-to-day basis.

3 See the Canadian Institute for Advanced Research, "The Pan-Canadian AI Strategy" (<https://cifar.ca/ai/>), which details strategy for supporting investment in AI development through universities and research institutions but does not identify the risks of AI use within research institutions.

# Definition of Terms

In this paper, “research institution” refers to institutions involved in the collection, storage, facilitation of access to, or dissemination of information and data in the interest of preserving and/or furthering human knowledge, history and culture, serving the public and social goods of research, education and innovation. This is primarily referring to public research libraries, archives and museums in Canada, but some relevant institutions within and outside of Canada are not public.

The definition of “AI” in this paper is based on the Organisation for Economic Co-operation and Development’s (OECD’s) definition of “a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.”<sup>4</sup> This definition addresses current and future AI that might be used in research institutions, including large language models (LLMs).

## Risk and Reward

The debate on how to legislate the use of AI has focused primarily on the struggle between maximizing value and benefit from AI tools and mitigating significant risks posed by these tools, including those that are deemed existential risks to humanity. There have been recent debates as to whether regulation should focus on the immediate risks of AI such as increased disinformation or on the “science-fiction-like” existential threats of artificial general intelligence (a broad term used to refer to AI systems that will outperform humans in most basic tasks). However, there are both immediate and future risks that must be addressed, and any policy framework for AI must be flexible enough to adapt to both eventualities. For research institutions, the policy must first and foremost preserve and enhance the role these institutions play in society — namely, the preservation and proliferation of useful and trustworthy data and information for the benefit of culture, science and the economy. This section will contextualize AI within the unique functions and needs of research institutions.

## Benefits

The primary benefits of AI use in research institutions are increasing access to and quality of data, improved labelling and contextualization of materials, and improving digitization and accessibility. The first potential benefit is to address the disparity between the high quality of metadata for library and archive catalogues of physical materials and the very low quality of metadata in online publications. Catalogues for physical collections receive extensive, expert human attention to ensure accuracy and depth in bibliographical descriptions. Digital resources tend to have their metadata generated without much attention paid to quality. For example, even the basic task of autogenerating a citation from a digital source often yields an incomplete citation that the user will need to correct before using it in a bibliography. AI-enabled catalogues, if trained well, would be able to learn what

---

<sup>4</sup> See <https://oecd.ai/en/ai-principles>.

high-quality data is and what is most useful to users, generating more relevant data over time. This function would allow for quick and accurate cataloguing of materials, which would solve the problem of cataloguing backup that every institution experiences, freeing time for other tasks such as outreach, teaching and analysis.

When there is high-quality and meaningful data about the author, the text, the argument, the publication details or the data used in a text, this enables an institution to generate linked data, which are points in common that can link two or more texts through keywords (see Heath and Bizer 2011). This capability exists now in a limited capacity. With AI tools, however, linked data could automatically link materials in ways that a human cataloguer might not be able to see when generating keywords, and it has the potential to generate entirely new areas of exploration across disciplines and methods. Transdisciplinary approaches are a major priority for institutions such as the Centre for International Governance Innovation's Digital Policy Hub.<sup>5</sup> Enhanced linking data would also give researchers the capacity to quickly generate bibliometric analysis, which is the practice of mapping citations and connections between different materials to understand the broader discourse and trends of a topic (see Linnenluecke, Marrone and Singh 2020). Bibliometric analysis is a work-intensive process that, with abundant metadata and AI tools, could become an automated function on a library's website.

High-quality metadata, bibliographical data, linking data and bibliometrics, would begin to produce a map of research that tracks thematic relationships over time: how texts change, are received and their perceived meaning in different periods, which could form the basis of context-specific labelling for research. In historical research, terminology and word meanings may change over time. With an AI-enabled search, historians could work with search terms that would exist relative to the context in which the searcher was exploring. For example, the term *historiografia* in sixteenth-century Portugal refers to a genre of narrative historical writing, while the modern translation, "historiography," refers to the practice of analyzing historical methods over time. The two terms are linked but have very different meanings that should be preserved in the interest of context, and the contextualization of the two meanings also creates opportunity for thinking more deeply about the purpose and nature of historical writing, which is always shifting. This is also relevant when thinking about materials from colonial or authoritarian contexts that use outdated and offensive language that should not be reflected in the basic catalogue but should still be searchable for historical research. For example, the Nazi label of "degenerate art" should not come up as a general keyword when one searches the painter Otto Dix. However, a historian searching for specific examples of so-called degenerate art should be able to find materials from the painters who were labelled as such, as well as contemporary works by those scholars who defined and supported the term (appropriately contextualized for the reader), and modern scholarly work that analyzes the historical policies and the discourse that shaped that term. This would also have value in understanding changing terms and discourse in scientific thinking, whether distinguishing between current views on the movement of planets and Nicolaus Copernicus's early writing on the topic or simply following the changing ideas in quantum mechanics over the last two decades. All of these uses allow for the historical record and context to be preserved while also supporting the need for a system that can adapt to the constantly changing methods and terms of research.

---

<sup>5</sup> See [www.cigionline.org/digital-policy-hub/](http://www.cigionline.org/digital-policy-hub/).

Ryan Cordell (2020, 519) discusses a version of changing the labelling of texts that he calls speculative bibliography, “in which textual associations are constituted propositionally, iteratively, and (sometimes) temporarily, as a result of probabilistic computational models.” A relative and potentially shifting labelling system would serve historians by using context-specific language that could change depending on the period or context in which a researcher was searching.

Speculative bibliography could also have a predictive capacity, to complement its historicizing and contextualizing function. It has been noted that generative artificial intelligence (GAI) can enhance research by identifying relationships in data that humans do not have the capacity to recognize (UNESCO 2023, 29). In the United States, the US National Science Foundation (2023) has announced a project where AI will be used to speculate on future paths for innovation and research, essentially delegating part of the research design and strategic planning process to an AI system that can plot the paths of greatest potential in scientific research. Having reliable and aligned AI systems for research in Canada could be the basis for similar predictive and speculative capabilities.

Finally, AI has a tremendous power to increase accessibility through improved optical character recognition technology that can scan physical texts to PDFs and enable accessibility features. Newer AI models have shown that this can also be done with images from older books, providing accurate descriptions, context and analysis of illustrations. This not only promotes accessibility for the visually impaired or those with other disabilities, but also for non-specialists looking at an unfamiliar image in an unfamiliar book (Weinberger 2023). Thinking broadly about the idea of accessibility, quick and high-quality digitization, contextualization of text and images, and the ever-improving ability for AI tools to process and translate between languages would also make different materials accessible across language barriers. This could enhance scientific research collaboration across Chinese, Spanish and English (for example), but it could also help to both preserve and share the complexities of writing in a specific language. There are ways of thinking that are bound to the logic of a language itself, such as the distinct way of knowing, which are integral to Indigenous languages and differ substantially from European languages. It is also important to note that colonization often included degradation of Indigenous languages, and a major element of reconciliation has been to reinvigorate these languages and promote their study. AI systems could help to support and expand these language revitalization efforts.

## Risks

The impact of AI in labour and job loss highlighted by the OECD is not a significant concern with research institutions. The potential uses of AI in research institutions would add capabilities that research institutions do not currently have or streamline time-consuming tasks that lead workers away from other important work. The more complex research, buying of materials, outreach and promotion of collections, teaching, advanced bibliographical analysis and institutional strategizing would become the primary focus of workers in research institutions, while the tasks of cataloguing, updating labels and managing loan services could be streamlined.

Information pollution is perhaps the largest risk posed by AI use in research institutions as all other issues of equity, access, preservation and reliability stem

from data quality. If the model used by a research institution is trained on poor-quality or biased data (which describes most currently available LLMs), then one of the fundamental duties of the institution — to preserve correct information and data — is undermined. A model that has incorrect information or faulty reasoning cannot be allowed to access what should be pristine data and would undermine the benefits of AI use in research institutions as they rely on high-quality data.

This connects to another of the OECD risk areas — the lack of diverse and quality data sets for AI training that would undermine the potential benefits and might lead to prejudice and discrimination or other biases. UNESCO highlights the risk of AI “reducing diversity of opinions and further marginalizing already marginalized voices” as a primary risk of AI in education and research contexts (UNESCO 2023, 17). Diversity must be defined broadly, including linguistic, geographical and cultural sources, but must also include a diversity of historical periods, research methods, values and any other diversity that reflects a pluralistic society. This is for the purpose of inclusion and equity but, perhaps more importantly, true diversity in data and in the people designing and implementing these tools will ensure any AI systems used in research institutions are of a robust nature, serving their purpose of supporting and enhancing human thought and research.

Regarding the risk of concentrating AI power in the hands of rich nations and the threat of malicious uses of AI, setting high standards for research institutions implementing AI systems will serve as a preventive measure. Having diverse and trustworthy publicly accessible research institution systems will help to promote access to information and to the benefits of AI. Having reliable stores of information in research institutions that are accessible and enabled with high-quality AI tools, will provide a deterrent to and possibly a means of combatting mis- and disinformation. However, poor-quality systems that lead to data pollution and engender distrust in institutions will contribute to the erosion of public trust.

The violation of norms (such as copyright and academic integrity) by AI tools is not a major concern with AI use in research institutions, as institutions and scholars are already bound by the law in these areas. However, the challenge of so-called hallucination (the creation of invented facts by LLMs) and the black-box nature of many models (the fact that the processes by which they arrive at conclusions or outputs are not fully knowable by users or developers) pose a threat both to following laws and to trust of models. As cognitive scientist Gary Marcus has written in multiple contexts, hallucinations are inherent to current LLMs and will never go away entirely (see Davis and Marcus 2023). The challenge of black-box models being used in research institutions is that the logic by which analysis or data is produced or organized might not be legible to users or administrators, making it difficult to rely on for research and to understand if it is aligned with human interests. Until models have eliminated hallucination and broken into the black box, so to speak, research institutions must consider limiting use of LLMs to ensure the quality of their digital systems.

Finally, the effects of AI on the environment are significant, but far from new for research institutions. Computer server banks and the physical spaces that support collections, students, researchers and staff consume high amounts of energy that must be curbed over time. This is already a concern for libraries and archives and will continue to be a concern regardless of how AI is deployed.



The listed risks are likely possibilities if AI were to be deployed in a research institution or network without effective policies in place. On a small scale, the loss of context, misleading results of a single hallucination or poor-quality metadata are inconveniences. However, when considering the centralized nature of research institutions, the volume of information in them, the number of critical projects that rely on them and that entire communities have their cultural heritage preserved in them, the risk of any large-scale degradation or misrepresentation becomes more serious. Just as AI has been called a threat to democracy (Kreps and Kriner 2023), the risk to research institutions is equally serious, endangering the fundamental structure of research and innovation on which our society relies.

## The Existing Legislative and Policy Landscape

It is important to distinguish between the proposed and existing regulations that govern AI research, meaning regulation that imposes restrictions and reporting standards onto developers making AI models that exceed certain thresholds of capacity or compute. These regulations and policies are specific to research in developing new AI, not the use of AI tools (that will presumably already have passed regulator and safety scrutiny) in research institutions. This paper aims to develop policy for how AI tools will manage sensitive information and data in research institutions, rather than regulating how research is conducted.

Thus far, the UNESCO AI research and education report is the only international document to propose any policy for research institutions, but it mostly focuses on the educational implications of AI, not on AI use within the internal systems of research institutions (UNESCO 2023, 5). The report calls for serious “open public debate and policy dialogues on the long-term implications of AI in education and research” including public and private stakeholders (ibid., 20). This paper is in part responding to the UNESCO call to action on AI in research, with the hope of building preliminary capacity to address AI risk. Some of the action items in the UNESCO report are: avoiding the erosion of human agency; overreliance on AI tools (ibid., 24-25), centring human creative outputs; and promoting a plurality of ideas (ibid., 25). This speaks to the cultural and social value of research institutions in supporting diverse opinions, critical thought and creativity.

UNESCO’s general AI development recommendations highlight that any AI legislation and policy is, and should be, rooted in existing human rights and international agreements (UNESCO 2022, 6). The necessary tools and frameworks for policy development already exist in many cases, and so it is important to not reinvent the wheel for what is perceived to be a paradigm-shifting technology, but to build on what exists and innovate in the gaps. Many national and international laws will automatically apply to AI in research institutions, especially regarding copyright, intellectual property, human rights and data. In the context of human rights, the most

applicable UN resolutions are the rights to culture,<sup>6</sup> access to science (UN General Assembly 2012) and the right to education (UNESCO 2019), which will influence how research institutions will implement AI. Beyond institutions, there is also the important discussion of how AI enablement in developed economies will affect developing economies, and the potential for an AI gap between countries such as Canada and those that do not have the same investment and research infrastructures.

Another useful international framework for developing policy for AI use in research institutions is the OECD's "AI Principles overview," which provides "values-based principles." These principles are: inclusive growth, sustainable development and well-being; human-centred values and fairness; transparency and explainability; robustness, security and safety; and accountability.<sup>7</sup> Based on these values, they also provide recommendations for policy makers: invest in AI research and development; foster a digital ecosystem for AI; provide an enabling policy environment for AI; building human capacity and preparing for labour market transition; and international cooperation for trustworthy AI. These are useful guides for thinking about the broader implications of an AI policy framework for Canadian research institutions and how it might be used as a base for significant international cooperation.

Within Canada, there is no legislation, proposed or existing, that addresses AI use in research institutions. The draft Canadian Artificial Intelligence and Data Act does not mention research at all and provides no specific guidance about the protection of valuable scientific, cultural and historical information in the research institution context.<sup>8</sup> The Ministry of Innovation, Science and Industry oversees this proposed legislation and therefore the development and use of technology is of primary concern, while implementation in research institutions does not seem to be a current focus. There are several other governmental bodies that might have an impact on AI policy in research institutions, but the most actively involved currently is the Office of the Chief Science Advisor, which has been in close dialogue with the Canadian Association of Research Libraries (CARL) on open science efforts and on early AI-related research. There is no direct legislation for AI in research institutions seemingly due to government resistance to regulate academic institutions and likely also due to a lack of recognition of the risks and benefits of AI use in research institutions.

## Other Relevant Policy Frameworks

Due to the lack of legislation and policy on AI use that is specific to research institutions, it is important to look to the institutions themselves for their AI-specific policies (which are, thus far, limited) and other relevant guidance that might be useful in building a sector-wide policy framework. Some useful building

---

6 UNESCO, *Convention Concerning the Protection of the World Cultural and Natural Heritage*, 16 November 1972, WHC/72/WS/1 (entered into force 17 December 1975), online: <<https://whc.unesco.org/en/conventiontext/>>.

7 See <https://oecd.ai/en/ai-principles>.

8 Bill C-27, *An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts*, 1st Sess, 44th Parl, 2022 (first reading 16 June 2022), online: <[www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading](http://www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading)>.

blocks from which to construct a policy for AI use within research institutions include the following: university policies for AI use; library and archive policies on ethics, data stewardship and social responsibility; and library and archive associations' policies for sustainable growth and technological advancement.

The first tier of university AI policies have very little formal policy and mostly provide students and faculty with basic resources for navigating the topic. The University of Ottawa,<sup>9</sup> the University of Regina,<sup>10</sup> the University of Manitoba<sup>11</sup> and the University of Saskatchewan<sup>12</sup> have guidelines or statements on ChatGPT, on AI use in the classroom, brief FAQ sections and links to resources. The University of British Columbia<sup>13</sup> also provides a map detailing when to use AI in the classroom and when to avoid it.

The second tier could be called the “inform and warn” policies. These focus on the risks of AI to academic integrity and copyright, and establish rules and guidelines for how students and professors can avoid these risks. Queen’s University’s policy focuses on privacy and academic integrity issues stemming from GAI.<sup>14</sup> Laval University’s policy discusses academic integrity but has no substantive policies yet.<sup>15</sup> Laval also has a statement for professors to include in their syllabi regarding GAI use in class (*ibid.*). The University of Waterloo has several documents focusing on issues of copyright violation, academic integrity and resources for students and professors.<sup>16</sup>

The third tier tends to present AI as a neutral tool that can be used and misused, stressing the need for education to ensure positive outcomes. Dalhousie University highlights the need for student and faculty education on AI tools and argues that professors should only use AI to enhance the learning experience rather than expediting a transactional process.<sup>17</sup> McGill University’s policy is based on five principles for GAI: AI education/literacy; the positive role of AI; professors’ discretion over AI use; high standards for instructors’ uses of AI; and student responsibility to existing rules.<sup>18</sup> The University of Calgary’s Taylor Institute for Teaching and Learning highlights how using AI is not inherently misconduct, that it does not threaten human creativity and that uses within classes should be integrated into the aims of the course (Eaton and Anselmo 2023). The University of Guelph has recommendations for addressing AI in the classroom focusing on professors’ discretion and integrating AI into courses.<sup>19</sup>

---

9 See <https://techlaw.uottawa.ca/aisociety/inclusion>; [www.uottawa.ca/research-innovation/impact-discoveries/artificial-intelligence](http://www.uottawa.ca/research-innovation/impact-discoveries/artificial-intelligence).

10 See <https://ctl.uregina.ca/generative-ai-at-the-university-of-regina>.

11 See <https://news.umanitoba.ca/academic-integrity-and-artificial-intelligence/>; <https://umanitoba.ca/centre-advancement-teaching-learning/um-syllabus-statements-genai>.

12 See <https://libguides.usask.ca/c.php?g=418130&p=5275298>.

13 See <https://academicintegrity.ubc.ca/generative-ai/>.

14 See [www.queensu.ca/ctl/resources/educational-technology/generative-ai-teaching-and-learning](http://www.queensu.ca/ctl/resources/educational-technology/generative-ai-teaching-and-learning).

15 See [www.ulaval.ca/en/artificial-intelligence](http://www.ulaval.ca/en/artificial-intelligence).

16 See [https://subjectguides.uwaterloo.ca/chatgpt\\_generative\\_ai/resources](https://subjectguides.uwaterloo.ca/chatgpt_generative_ai/resources); <https://uwaterloo.ca/copyright-at-waterloo/teaching/generative-artificial-intelligence>; <https://uwaterloo.ca/writing-and-communication-centre/Resources-AI-Overview>; [https://subjectguides.uwaterloo.ca/chatgpt\\_generative\\_ai/citationimportance](https://subjectguides.uwaterloo.ca/chatgpt_generative_ai/citationimportance).

17 See [www.dal.ca/dept/ctl/e-learning/AI\\_Resource/a-i--and-academic-integrity.html](http://www.dal.ca/dept/ctl/e-learning/AI_Resource/a-i--and-academic-integrity.html); [www.dal.ca/dept/ctl/e-learning/AI\\_Resource.html](http://www.dal.ca/dept/ctl/e-learning/AI_Resource.html); [www.dal.ca/dept/ctl/e-learning/AI\\_Resource/designing-assessments-with-a-i--in-mind.html](http://www.dal.ca/dept/ctl/e-learning/AI_Resource/designing-assessments-with-a-i--in-mind.html); [www.dal.ca/faculty/open/about/faculty-statement-on-generative-artificial-intelligence.html](http://www.dal.ca/faculty/open/about/faculty-statement-on-generative-artificial-intelligence.html).

18 See [www.mcgill.ca/stl/stl-ai-working-group](http://www.mcgill.ca/stl/stl-ai-working-group); McGill University (2023).

19 See <https://otl.uoguelph.ca/teaching-assessment-resources/teaching-context-ai/provisional-recommendations-use-generative-ai>.

The fourth tier of university AI policies consists of institutions that have departments dedicated to AI policy and implementation. McMaster University has a Task Force on Generative AI in Teaching and Learning, which guides their policy development.<sup>20</sup> As a result, McMaster has a list of principles for GAI use in the university<sup>21</sup> and has general guidelines for professors using AI in teaching or in research, highlighting the importance of “Privacy, Security and Selection of Tools.”<sup>22</sup> The University of Alberta’s Provost’s Taskforce on Artificial Intelligence and the Learning Environment has not yet produced a major policy document.<sup>23</sup>

The University of Montreal has a heavy focus on AI ethics, likely due to the involvement of the AI scientist and advocate for AI regulation, Yoshua Bengio, who is part of the university’s AI steering committee and played a role in its responsible development declaration.<sup>24</sup> The steering committee produced a statement guiding AI research that includes 10 principles for AI development at the university. The principles are: well-being; autonomy; privacy; solidarity; democracy; equity, diversity and inclusion; prudence; responsibility; and sustainability.<sup>25</sup> The statement also lists recommendations for development that highlights international collaboration.

The University of Toronto does not currently have a distinct AI office, but with several AI institutes at the university, there are a variety of bodies engaged in AI policy. One of the university’s more unique documents is its “Guidance on the appropriate use of Generative Artificial Intelligence in Graduate Theses,” which applies their policies to the doctoral thesis process in a practical and clear way.<sup>26</sup> The University of Toronto also has documents on academic integrity and AI, AI tool use by instructors<sup>27</sup> and suggested language for syllabi (University of Toronto 2023).

Western University is the first school in Canada to have a chief AI officer, to navigate the changing AI landscape (Ferguson 2023). The university also has written guidance for how individual roles in the university should approach AI use, including for researchers, mostly highlighting existing commitments and policies.<sup>28</sup> When comparing Canadian university AI policies with those from other nations, Canada is a clear leader in AI policy in higher education (see Usher and Desforges 2023). This may be due to the central role Canadian institutes and universities played in the AI boom worldwide. All policies by non-Canadian universities follow the trends in the first tier of Canadian university policies.

---

20 See <https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/>.

21 See <https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/>.

22 See <https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/>.

23 See [www.ualberta.ca/provost/policies-and-procedures/taskforce-on-artificial-intelligence-and-the-learning-environment.html](http://www.ualberta.ca/provost/policies-and-procedures/taskforce-on-artificial-intelligence-and-the-learning-environment.html).

24 See [www.umontreal.ca/en/artificialintelligence](http://www.umontreal.ca/en/artificialintelligence).

25 See <https://montrealdeclaration-responsibleai.com/>.

26 See [www.sgs.utoronto.ca/about/guidance-on-the-use-of-generative-artificial-intelligence/](http://www.sgs.utoronto.ca/about/guidance-on-the-use-of-generative-artificial-intelligence/).

27 See [www.viceprovostundergrad.utoronto.ca/strategic-priorities/digital-learning/special-initiative-artificial-intelligence/](http://www.viceprovostundergrad.utoronto.ca/strategic-priorities/digital-learning/special-initiative-artificial-intelligence/); [www.academicintegrity.utoronto.ca/perils-and-pitfalls/using-chatgpt-or-other-ai-tool-on-a-marked-assessment/](http://www.academicintegrity.utoronto.ca/perils-and-pitfalls/using-chatgpt-or-other-ai-tool-on-a-marked-assessment/).

28 See <https://ai.uwo.ca/>.

Above, we see how universities have tried to anticipate AI's effects and put into place some international frameworks for mitigating AI risk. In the absence of a specific framework for how research institutions should use AI, the existing policies and codes of ethics from associations of library and archive workers are the best body for building a sense of the shared fundamental values and goals of these institutions. Many of these institutions are actively researching policy options for AI use, but there is not yet any specific framework and there is no inter-institutional collaboration on policy. Each institution has unique goals and practices based on the kinds of materials they use and the publics they serve. However, all these institutions have a shared purpose to preserve and disseminate knowledge in society and can be linked to a set of core principles that intersect all research institutions.

It is also important to call attention to the existing algorithms and machine-learning elements already used for decades in library catalogues and research institutions, which constitute forms of AI and should be governed by robust policy. These existing uses are primarily in library catalogue searches and metadata organization within the institutions but tend to be based on proprietary algorithms owned by the companies selling the library management service that a library uses. The sector already has decades of experience adapting new technologies to serve their function as the historical memory and record of society and are thus highly qualified to think about how to align AI with their goals.<sup>29</sup>

The Association of College and Research Libraries (ACRL) Code of Ethics for Special Collections Librarians is the standard for most special or rare collections in Canada and the United States. The code holds practitioners to a high standard “due to the significant cultural and monetary value of the materials held in special collections libraries” and the fact that they preserve “cultural property, preserving original artifacts, and supporting both community engagement and scholarship with primary research materials” (ACRL 2020). The ACRL code is based on the core values of access and accessibility, description, development, labour practices and professionalism, outreach, reference, instruction and exhibitions, preservation and security, privacy and responsibility to originators of materials (ibid.). These values focus primarily on cultivating and preserving good quality materials for collections, protecting materials and ensuring access to them is safe and accessible to society.

The Society of American Archivists (SAA) argues that archivists should: expand access; contribute to scholarship; collaborate with users; be transparent; respect diversity and ensure archives reflect that; pursue professional development; preserve and share collections sustainably; mentor newcomers to the field; and share knowledge with users.<sup>30</sup> These principles align with the ACRL code with the addition of the obligation to mentor new archivists in the field. They summarize their purpose and values as: preserving the cultural heritage of society through documentary record; compiling and maintaining records at all levels of society; “assisting in the interpretation of documentation of past events through the use of primary source materials”; and serving

---

29 In the introduction to *The Rise of AI: Implications and Applications of Artificial Intelligence in Academic Libraries*, the editors argue that “Librarians are uniquely positioned to rise to the challenge that AI presents to their field. Libraries and their like have existed for millennia; they progress with society, altering and adapting their services to meet the information needs of their communities” (Hervieux and Wheatley 2022).

30 See [www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics](http://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics).

all people interested in using the documents.<sup>31</sup> The SAA code of ethics encourages professional relationships, good judgment (transparency about curation of materials), authenticity (origins of materials), protection, access (within acceptable limits imposed by the collection), privacy (for users) and trust.<sup>32</sup> The ethical codes of both the ACRL and SAA maintain a high burden of responsibility to preserve and protect materials under the care of librarians and archivists, which are sufficiently robust to be applied to AI use in research institutions.<sup>33</sup> These existing ethical codes, followed by most Canadian and US research institutions, demonstrate that many of the necessary tools are already in place for sector-wide policy, given the proper coordination and support.

CARL has similar values and goals to its US counterparts, but with the added priorities of advancing research, advancing teaching and learning, strengthening capacity, measuring impact and influencing policy according to the topic sections on their website.<sup>34</sup> CARL's distinct policy role places it in a position to lead Canada-wide and international policy development efforts for AI use across research institutions. Another priority for CARL that differs from its counterparts is protecting Indigenous and other traditional knowledge. CARL commits to balancing respect for tradition with open access; seeking to maximize the latter if it is in keeping with the former, following the International Federation of Library Associations and Institutions' (IFLA's) guiding principles on traditional cultural expressions and knowledge.<sup>35</sup> The precarity of Indigenous knowledge is an important priority for research institutions in Canada and should be central to any discussion of equity and diversity in AI policy for research institutions. Indigenous materials are likely to be in collections under problematic arrangements, and often these collections constitute the entirety of written record of a community. Such materials should only be accessed by trustworthy AI with the community's informed consent with reference to other frameworks such as OCAP®.<sup>36</sup>

CARL's Strategic Focus 2022-2025 declares a plan to build leadership on behalf of research libraries and develop local, national and international collaboration focused on digital culture, changing dynamics of research, technological development and increasing interest in equity to enhance Canada's existing leadership role in the field (CARL 2022). These strategic priorities take the principles shared by existing associations and put them in the context of future-oriented policy development to address technological changes in the sector and society. As part of this concerted effort to develop and shape policy for research libraries, CARL has also submitted several policy papers to the Government of Canada, including a recommendation that the Government of Canada invest significant funds to support an "AI and Data

---

31 Ibid.

32 Ibid.

33 See related policies from the Association of Canadian Archivists ([https://archivists.ca/ACA\\_strategic\\_framework\\_2023-2026](https://archivists.ca/ACA_strategic_framework_2023-2026); <https://archivists.ca/Reconciliation-Framework>), the Association of Research Libraries ([www.arl.org/who-we-are/#section-codeofconduct](http://www.arl.org/who-we-are/#section-codeofconduct); [www.arl.org/who-we-are/](http://www.arl.org/who-we-are/)) and the Canadian Council of Archives (<https://archivescanada.ca/what-we-do/>).

34 See [www.carl-abrc.ca/](http://www.carl-abrc.ca/).

35 See [www.carl-abrc.ca/influencing-policy/traditional-knowledge-2/](http://www.carl-abrc.ca/influencing-policy/traditional-knowledge-2/); IFLA (2012).

36 OCAP® are First Nations principles for information governance, which stands for: Ownership: communities collectively own their information, and individuals own their own information; Control: First Nations have the right to control their information and the research processes that pertain to them; Access: communities have access to their own data and information wherever it is held and the right to decide who accesses that data and how; and Possession: stewardship over a community's information whether it is held by the community or on their behalf; see <https://fnigc.ca/ocap-training/>.

Commissioner...[to] oversee activities and respond to issues expediently” (CARL 2023c, 5). CARL’s argument on the “Right to Repair, Interoperability and Technical Protection Measures,” proposed to the Government of Canada in a report with the Canadian Federation of Library Associations (CFLA), promotes institutional collaboration and standardization of technology for a resilient and useful research network in Canada and internationally (CFLA and CARL 2021, 9). CARL’s recent work, in combination with the ACRL and SAA frameworks, represent the sector-wide ethical and practical norms and standards from which one can effectively extrapolate key issues for AI governance.

## Proposed Framework

Based on the shared responsibilities, values and obligations of research institutions, and on the primary concerns of the national and international discussion on AI governance, a foundational governance framework for AI use in research institutions should consist of five core values:

- **Preserving the integrity of information and data:** Protect and preserve physical and digital texts and materials of all kinds from censorship, arbitrary change, loss or removal from collections. This includes keeping a secure record of all materials held in a catalogue, and their relevant metadata, in case of system failure or data loss.
- **Preserving context:** Collect and preserve relevant and accurate bibliographical data, provenance, historical contextual data, appropriate and accurate keywords and labels, and other relevant metadata.
- **Promoting informed and principled accessibility:** This refers to digitization and other tools to support accessibility for disabled people and the public (not limited by class, culture or geography), to promote open science and access to education and to do so with informed consent from the owners or stakeholders of the information. This includes limiting access, when appropriate, in cases of sensitive data and to respect specific community interests (including Indigenous data governance). The tendency should be toward accessibility while respecting the wishes of relevant community and governing bodies.
- **Sustainability:** Defined broadly to include sustainable use of materials and energy, sustainable practices in ensuring the longevity of collections, and sustainability in the agreements and relationships governing data, prioritizing long-term access and relationships.
- **Transparency:** Ensure transparent labelling of information that is AI-generated or modified, openness to users about how machine learning is integrated into information systems and acknowledgement of potential inaccuracy of machine-generated materials.<sup>37</sup>

---

<sup>37</sup> Cordell’s (2022, 140–41) article “Closing the Loop” explores how machine learning should be implemented in research libraries, suggesting that transparently highlighting the advantages and limitations of machine learning could demystify the technology within research libraries, making its advantages and disadvantages clear, and could also be used to invite feedback from users to improve automated research systems.

This framework is simplified to the most basic shared values and priorities to ensure interoperability and relevance across research institutions and communities. It should serve as a basis for discussion across leading research institutions, associations and government to build consensus on the issue. Once these or other principles are agreed upon, institutions should develop policies that are specific enough to be effective but broad enough to be flexible to future challenges. Policies should be evidence-based and binding, with some means of compelling institutions to adhere to the policies. Institutions should develop a working group and/or an institutional body that will continue to implement, monitor, research and develop policy and tools to address AI across the sector. This might be a sub-department of a government body for AI regulation, an inter-institutional body hosted by one existing independent organization such as CARL, or a new independent body. Canadian institutions should establish a dialogue and promote similar policies at an international level to first address spaces where research networks overlap significantly (with the United States, the United Kingdom, the European Union, among others) and then look to establish consensus internationally. This is necessary to ensure safety and consistency across institutions, but also to promote the benefits of these tools in places where they may not otherwise be able to benefit, such as the developing economies.

## Conclusion

Research institutions form an as-yet-unregulated sector of society on which so many of our economic, political, scientific and cultural systems rely. Institutions in this sector will inevitably adopt some AI tools to enhance and streamline their various functions as tools emerge. Because of the foundational importance of these institutions and the vast amount of information they hold, AI use in the sector could have catastrophic results if implemented poorly or with the wrong tools, leading to extensive losses or the pollution of essential human knowledge, undermining other sectors and institutions in the process. Thus, it is essential that institutions and governments take decisive action on the issue, develop policy to address the risks and ensure the potential benefits of AI tool deployment across the sector. Much of the necessary policy and practices already exist in some form within these institutions but have not been implemented in a cross-institutional way to govern AI use in research institutions. This places Canada in a position to demonstrate leadership on the world stage in building international governance and capacity for AI-enabled research networks.

## Acknowledgements

The author would like to thank Paolo Granata, David Fernández, Aaron Shull and Paul Samson.



---

## About the Author

Matthew da Mota is a post-doctoral fellow with the Digital Policy Hub, researching the uses and governance of artificial intelligence and machine learning within universities and research institutions (including research libraries and archives). He is also exploring the implication of technology on research security, innovation policy, infrastructure and media. Matthew's other research explores the connections between information and power, historical writing within empires, the philosophy of history, propaganda and disinformation, and the way that media and technologies shape the circulation and control of information. He recently obtained his Ph.D. from the University of Toronto's comparative literature program.

# List of Acronyms and Abbreviations

ACRL	Association of College and Research Libraries
AI	artificial intelligence
CARL	Canadian Association of Research Libraries
CFLA	Canadian Federation of Library Associations
GAI	generative artificial intelligence
LLMs	large language models
OCAP®	Ownership, Control, Access and Possession
OECD	Organisation for Economic Co-operation and Development
SAA	Society of American Archivists
U15 Canada	U15
UNESCO	United Nations Educational, Scientific and Cultural Organization

# Works Cited

- ACRL. 2020. "Code of Ethics for Special Collections Librarians." American Library Association. Rare Books and Manuscripts Section. October. [https://rbms.info/standards/code\\_of\\_ethics/](https://rbms.info/standards/code_of_ethics/).
- CARL. 2022. "CARL's Strategic Focus 2022-2025." *Canadian Association of Research Libraries* (blog). [www.carl-abrc.ca/about-carl/carls-strategic-focus-2022-2025/](http://www.carl-abrc.ca/about-carl/carls-strategic-focus-2022-2025/).
- — —. 2023. "2024 Federal Budget: CARL Brief to House of Commons' Standing Committee on Finance." August. Ottawa, ON: CARL. [www.carl-abrc.ca/wp-content/uploads/2023/08/CARL-Brief-to-House-of-Commons-Standing-Committee-on-Finance\\_2024-Federal-Budget\\_Final.pdf](http://www.carl-abrc.ca/wp-content/uploads/2023/08/CARL-Brief-to-House-of-Commons-Standing-Committee-on-Finance_2024-Federal-Budget_Final.pdf).
- CFLA and CARL. 2021. "Brief to the Government of Canada Consultation on a Modern Framework for Artificial Intelligence and the Internet of Things." Submitted by Canadian Federation of Library Associations and Canadian Association of Research Libraries. September.
- Cordell, Ryan. 2020. "Speculative Bibliography." *Anglia* 138 (3): 519–31. <https://doi.org/10.1515/ang-2020-0041>.
- — —. 2022. "Closing the Loop: Bridging Machine Learning (ML) Research and Library Systems." *Library Trends* 71 (1): 132–43. <https://doi.org/10.1353/lib.2023.0008>.
- Davis, Ernest and Gary Marcus. 2023. "Hello, Multimodal Hallucinations." *Marcus on AI* (blog), October 21. <https://garymarcus.substack.com/p/hello-multimodal-hallucinations>.
- Eaton, Sarah and Lorelei Anselmo. 2023. "Teaching and Learning with Artificial Intelligence Apps." Taylor Institute for Teaching and Learning, University of Calgary. January. <https://taylorinstitute.ucalgary.ca/teaching-with-AI-apps>.
- European Commission. 2023. "Regulatory Framework Proposal on Artificial Intelligence: Shaping Europe's Digital Future." *Shaping Europe's digital future*, European Union, November. <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>.
- Ferguson, Keri. 2023. "Western appoints Mark Daley as first-ever chief AI officer." *Western News*, September 27. <https://news.westernu.ca/2023/09/western-appoints-daley-chief-ai-officer/>.
- Heath, Tom and Christian Bizer. 2011. *Linked Data: Evolving the Web into a Global Data Space*. 1st ed. Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. San Rafael, CA: Morgan & Claypool.
- Hervieux, Sandy and Amanda Wheatley, eds. 2022. *The Rise of AI: Implications and Applications of Artificial Intelligence in Academic Libraries*. Chicago, IL: ACRL.
- IFLA. 2012. "Guiding Principles for IFLA's position concerning international treaties relating to traditional cultural expressions." March 23. [www.ifla.org/publications/guiding-principles-for-iflas-position-concerning-international-treaties-relating-to-traditional-cultural-expressions-2012/](http://www.ifla.org/publications/guiding-principles-for-iflas-position-concerning-international-treaties-relating-to-traditional-cultural-expressions-2012/).
- Kreps, Sarah and Doug Kriner. 2023. "How AI Threatens Democracy." *Journal of Democracy* 34 (4): 122–31. [www.journalofdemocracy.org/articles/how-ai-threatens-democracy/](http://www.journalofdemocracy.org/articles/how-ai-threatens-democracy/).
- Linnenluecke, Martina K., Mauricio Marrone and Abhay K. Singh. 2020. "Conducting systematic literature reviews and bibliometric analyses." *Australian Journal of Management* 45 (2): 175–94. <https://doi.org/10.1177/0312896219877678>.
- McGill University. 2023. "STL approves recommendations on generative AI." Channels: McGill University News, July 26. [www.mcgill.ca/channels/channels/news/stl-approves-recommendations-generative-ai-349064](http://www.mcgill.ca/channels/channels/news/stl-approves-recommendations-generative-ai-349064).

- U15. 2023. "U15 Canada Provides Guidance on Leading Institutional Practices to Safeguard Research." U15 Group of Canadian Research Universities. <https://u15.ca/publications/statements-releases/u15-canada-provides-guidance-on-leading-institutional-practices-to-safeguard-research/>.
- UNESCO. 2019. *Right to education handbook*. Paris, France: UNESCO and Right to Education Initiative. <https://doi.org/10.54675/ZMNJ2648>.
- – –. 2022. "Recommendation on the Ethics of Artificial Intelligence." May. Paris, France: UNESCO. [www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence](http://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence).
- – –. 2023. "Guidance for generative AI in education and research." September 7. Paris, France: UNESCO. [www.unesco.org/en/articles/guidance-generative-ai-education-and-research](http://www.unesco.org/en/articles/guidance-generative-ai-education-and-research).
- UN General Assembly. 2012. "Report of the Special Rapporteur in the field of cultural rights, Farida Shaheed: The right to enjoy the benefits of scientific progress and its applications." A/HRC/20/26. May 14. <https://digitallibrary.un.org/record/730844?ln=en>.
- University of Toronto. 2023. "U of T Syllabus Language – Use of Generative AI in Assignments." Office of the Vice-Provost, Innovations in Undergraduate Education. April. [www.viceprovostundergrad.utoronto.ca/wp-content/uploads/sites/275/2023/04/Syllabus-Language-for-Gen-AI-April-2023.pdf](http://www.viceprovostundergrad.utoronto.ca/wp-content/uploads/sites/275/2023/04/Syllabus-Language-for-Gen-AI-April-2023.pdf).
- US National Science Foundation. 2023. "New report identifies pathways to strengthen U.S. competitiveness in key technology areas." US National Science Foundation Directorate for Technology, Innovation and Partnerships update, October 24. <https://new.nsf.gov/tip/updates/new-report-identifies-pathways-strengthen-us-competitiveness>.
- Usher, Alex and Sandrine Desforges. 2023. "AI in Canadian Higher Education – An Update." Higher Education Strategy Associates, October 18. <https://higherstrategy.com/ai-in-canadian-higher-education-an-update/>.
- Weinberger, Adam. 2023. "AI ChatGPT and Rare Books." October. YouTube video, 17:02. [www.youtube.com/watch?v=SHAU5BnRMC8](http://www.youtube.com/watch?v=SHAU5BnRMC8).