Centre for International Governance Innovation

Digital Policy Hub

Digital Policy Hub – Working Paper

AI Governance Needs a Climate Change Strategy

Christelle Tessono

Winter 2024 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.

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.

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.

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

Key Points

- Large language model (LLM) powering systems, such as ChatGPT and Meta AI, are rapidly evolving technologies, and their adoption has become increasingly fast paced.
- However, legislation has been unable to catch up to speed. In Canada, the proposed Artificial Intelligence and Data Act (AIDA) has been sitting idly since June 2022 with very limited parliamentary traction, but with significant multi-sectoral criticism due to a lack of public consultation.
- With increasing investments in Canada's artificial intelligence (AI) computing capacity, little attention has been paid to the growing environmental harms that emerge from scaled-up AI developments.
- To address those harms, this working paper proposes including environmental impact assessments and corresponding limitations on environmentally degrading developments.

Introduction

It took Twitter 65 months post-launch to gain 100 million users; it took OpenAI's ChatGPT only two months to achieve that same milestone (Leslie and Rossi 2023). In light of this rapidly growing market amid consumer interest for generative AI, the federal government has responded by developing voluntary codes and guidelines to ensure responsible development and use of these technologies (Lu 2023). The establishment of these codes and guidelines has been met with criticism as they are not substitutes for the robust legislative protection of human rights. But with increased market incentive for development and adoption, there are a multitude of considerations that policy makers have to take into account beyond human rights — most notably, the environmental impact of generative AI. An increasing number of studies are reporting that LLMs, the form of generative AI behind ChatGPT, are contributing to climate change due to their intense resource extraction activities (Bender et al. 2021; Luccioni, Viguier and Ligozat 2023). Considering the multiplicity and unevenness of climate change ramifications, action needs to be taken to ensure that technological developments do not contribute to environmental collapse. This working paper provides a brief discussion of how LLMs impact the environment as well as the limitations of Canada's proposed AIDA in response to these emerging environmental harms. Mitigation measures are needed across the lifecycle of these technologies to effectively address and prevent direct and indirect harms caused by the rapid development and deployment of LLMs, and the conclusion of this paper provides key avenues that can be readily adopted by the federal government.

What Are LLMs and How Do They Work?

LLMs are a product of natural language processing (NLP), an interdisciplinary field emerging at the intersection of computer science, linguistics and statistics. NLP falls under the umbrella of AI, with one of its primary goals being to build computerized systems that can process human language in a diversity of forms (such as data sets, books, websites and so on). A language model can be defined as "any system trained only on the task of string prediction," meaning that such a system is only able to generate text based on content on which it has been trained (Bender and Koller 2020, 5158). At a technical level, the task of LLMs is to accurately predict the likelihood of properly using the right words in a specific context; they cannot interpret language in the same way that the human brain can.

A great number of applications are powered by LLMs, notably Meta's LLaMA, Google's PaLM and Open AI's GPT-4. LLMs can be used to complete a great number of tasks, which include (but are not limited to) drafting, editing, translating and summarizing texts. They can also be applied in the generation of coding language and in automated audio transcription. As a result, LLMs have been deployed across different sectors to perform a variety of applied tasks. In health care, for example, LLMs are being used as "chronic-care nurses" as well as for notetaking and clinical diagnosis (Reese et al. 2024; Webster 2023); in the financial sector, they are being used for market prediction, algorithmic trading and risk assessment (Hadi et al. 2023).

LLMs' ability to accurately produce content and effectively perform tasks is the subject of great academic, industry and governance discussion. There are an increasing number of reports outlining this rapidly evolving technology's inaccuracies, limitations, privacy risks and socio-economic harms. However, consensus has yet to be achieved on a variety of factors such as standardization and model capabilities (Bender et al. 2021; Solaiman et al. 2023). For example, researchers have demonstrated the lack of robustness in these systems and have been disputing their emergent capabilities by providing evidence when employing new measurement methods (Schaeffer, Miranda and Koyejo 2023). Bad actors can also control LLMs by indirectly injecting prompts that compromise the operating system, thus allowing them to access and retrieve information such as personal data (Greshake et al. 2023).

Developments in the past decade have created more sophisticated LLMs due to investments in computing power and access to large amounts of data on which to train models (Leslie and Rossi 2023). In Canada, we have seen calls to increase the country's computing infrastructure from industry and academic actors (Castaldo 2024). Notably, when testifying in the Bill C-27 committee study, the Canadian Institute for Advanced Research advocated for increased funding to build computing power to support the country's growing AI ecosystem (Strome 2024). Compared to countries such as the United States and Japan, Canada lacks access to AI computing infrastructure despite being a hub for research and tech start-ups (Dobbs and Hirsch-Allen 2024). As a result, there are concerns that the lack of domestic computing infrastructure will lead to a decline in growth, productivity and innovation. The federal

government answered these concerns by proposing to invest $2 billion into computing infrastructure as part of the 2024 federal budget (Office of the Prime Minister 2024).

There is a growing body of scholarship dedicated to identifying the environmental implications behind the design, development and deployment of LLMs. Carbon emissions from an LLM come both from the initial build of the model (such as manufacturing the hardware components) and its continued operations (such as running the model and data centres) (Smith et al. 2023). Researchers have identified that while some computing providers' energy comes from renewable energy or carbon credit–offsetting sources, a majority of it does not (Bender et al. 2021; Edwards, Cooper and Hogan 2024). Moreover, renewable energy does not mean there is no harmful impact on the environment. To build wind farms, for example, whole ecosystems have to be destroyed, as was the case in Scotland when 14 million trees were cut down to make way for wind farms (Bol 2020). This has pushed researchers in the field to call for the development of even more energy efficient models (Bender et al. 2021). There is also concern about the amount of water consumption behind the operation of these tools: for example, researchers recently found that ChatGPT consumes 500 mL of water in responding to queries (Li et al. 2023). In addition, in a recent environmental report, Microsoft disclosed a 34 percent increase in their global water consumption between 2021 and 2022 due to power generative AI innovation (Kahil 2023).

One of the biggest challenges when it comes to the governance of LLMs consists of identifying the necessary policy interventions along the complex supply chains that make up these technologies. Life cycles are hard to draw out as each AI system may have "a different supply chain, with variations depending on the sector, the use case, whether the system is developed in-house or procured, and how the system is made available to those who use it" (Brown 2023, 6). To develop robust AI legislation, policy makers will need to "codify multiparty, gapless, and end-to-end accountability and transparency mechanisms" that facilitate the establishment of "a continuous chain of human responsibility" throughout the life cycle of systems in scope (Leslie and Rossi 2023, 3).

What Is the Policy Landscape for LLMs in Canada?

The Canadian AI policy landscape consists of a patchwork of strategies, directives, localized initiatives and proposed legislation (Attard-Frost, Brandusescu and Lyons 2023; Attard-Frost and Hayes 2023). This has led to an ecosystem that is decentralized and lacks uniformity across the private and public sectors. More specifically, the federal public sector does not have a statutory framework governing the use of AI systems by departments and agencies. Instead, it has the Directive on Automated Decision-Making (DADM), which is intended "to ensure that automated decision systems are deployed in a manner that reduces risks to clients, federal institutions and Canadian society" by establishing transparency, reporting, auditing and assessment requirements.[1] However, there are significant drawbacks to the DADM: it cannot be enforced as it is not legally binding legislation and it also excludes the internal activities of departments,

1 See www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592.

such as the use of AI for making hiring decisions (Toronto Metropolitan University, The Dais and the Centre for Media, Technology and Democracy 2023). Moreover, the DADM does not provide recourse, which makes it ill-equipped to address human rights and public safety concerns related to AI deployment (Tessono et al. 2022).

The proposed AIDA, which is part of Bill C-27, also has a lot of limitations. Presented in June 2022, the AIDA seeks to regulate the private sector trade and commerce of AI systems by requiring companies to adopt mitigation measures and transparency reporting practices.[2] In November 2023, the minister of innovation, science and industry, François-Philippe Champagne, tabled amendments in committee to clarify the role of the proposed AIDA commissioner, include a schedule for systems that would be considered in scope, and add special provisions relating to machine-learning models and general-purpose systems (Champagne 2023). With its weak human rights protections, and relegation of major components to regulations, plus a lack of public consultation, the AIDA has received significant pushback from civil society, industry and academia (Tessono et al. 2022).

There are two government initiatives specific to generative AI: the Treasury Board Secretariat's "Guide on the Use of Generative Artificial Intelligence" and the Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems. The guide was published in 2023 with the intention of providing departments and agencies guidance on how to responsibly and safely use generative AI tools.[3] In September 2023, the federal government launched its Voluntary Code of Conduct, which invites developers and managers of generative AI systems to commit to implementing measures that will ensure safety, transparency, accountability, validity, robustness, fairness, equity and responsible human oversight of their technologies (Innovation, Science and Economic Development Canada 2023). While 23 companies have signed this agreement at the time of this writing, the code lacks specificity on how to achieve the measure it outlines and is unfortunately not enforceable by law.

As a result, there are four sets of important policy gaps that need to be addressed in the federal context: the limitations of Bill C-27; the absence of enforceable public sector legislation addressing AI systems; the absence of strong legislative provisions related to generative AI systems; and, lastly, the exclusion of environmental protections in AI development and deployment. The uneven application of AI governance facilitates "problematic double-standards" and "further distance[s] Canada from global accountability measures" (Tessono et al. 2022, 9).

2 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) [*AIDA*], online: <www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading>.

3 See www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html.

Harms along the Life Cycle Need to Be within AI Governance's Scope

In light of the growing concerns related to generative AI systems, Champagne introduced amendments that would add provisions related to general-purpose systems, including impact assessments, risk mitigation measures, mandatory reporting of incidents and record-keeping (Champagne 2023). However, these impact assessments would only relate to the narrow scope of harms, which the bill defines as "physical, psychological, damage to an individual's property or economic loss."[4] Moreover, these impact assessments would only focus on harms related to "biased output," which consists of content that "adversely differentiates, directly or indirectly and without justification" on the prohibited grounds of discrimination as outlined in the Canadian Human Rights Act (ibid.). As a result, the minister's proposed amendments to the AIDA fall short because they only require companies to demonstrate compliance with biased output and a limited set of harms. Not included in the scope of this legislation are different types of harm that can occur along the life cycle of LLMs, such as the environmental implications of these technologies as discussed earlier. It is important to consider the environmental implications of LLMs given the uneven distribution of catastrophic events caused by global warming. We are witnessing in real time the intensification of these events with summer 2023's devastating wildfires displacing entire communities and polluting the air across the country and into the United States. It is expected that in the future, wildfire disasters will not only become more frequent, but also more severe, leading to greater physical and ecological losses (Erni et al. 2024). As a result, taking action now on factors that contribute to climate change — including AI, and specifically LLMs — is imperative. While it is understandable that organizations engaged in the development of AI systems have "expressed concerns about being held accountable for post-deployment responsibilities" regarding the development of AI systems, it is nonetheless important that all actors involved be held accountable for harms that may occur during both the design and development stages (Champagne 2023, 7).

As argued by scholars Blair Attard-Frost and Helen Hayes (2023, 12), "the AIDA and other Canadian initiatives for AI value chain governance should widen the scope of actors and resourcing activities." This could be achieved by expanding governance to factors beyond data collection and output generation to the environmental implications present throughout the AI life cycle (ibid.). The AIDA should instead include specific provisions related to manufacturing, e-waste disposal and transport recurring throughout the life cycle of LLMs, to name a few examples.

4 See *AIDA*, *supra* note 2, s 5.1.

Recommendations

To develop a governance framework that is agile, robust and capable of withstanding rapid technological innovation in the AI space, the Canadian government must consider harms that emerge throughout the life cycle of AI technologies such as LLMs, including environmental harms. To move forward, the following three measures should be adopted:

- First, the AIDA should include provisions requiring companies to conduct environmental impact assessments. These assessments could include identifying how design choices, hardware, data centre location, model size and type contribute to environmental degradation and accelerate climate change (Laranjeira de Pereira 2024). As with other impact assessments proposed in the legislation, details would be prescribed in the regulations so that they would be adaptable to the pace of technological innovation in the AI space.
- Second, the government should develop corresponding limitations on environmentally degrading infrastructural developments and activities. To do this, policy makers across departments and agencies need to work together, alongside industry actors and members of both academia and civic society, to identify which design, development and operating choices contribute to climate change–accelerating activities such as water extraction. This aim could be achieved by fostering cross-departmental initiatives and hiring subject-matter experts across different fields.
- Third, the AIDA should be amended to consider both private and public sector institutions as part of the legislation's scope. Canada's existing AI governance strategy consists of a decentralized and unenforceable patchwork of AI policies, which create an uneven protection against harms. Harmonizing approaches would ensure accountable development, deployment and adoption across sectors. As a fallback, the government should commit to developing public sector–specific AI governance legislation.

Acknowledgements

The author would like to thank Ivan Nuñez Gamez, Amrita Vasudevan, Shion Guha, Helen Hayes, Mohammed Odusanya, Tristan Masson, Christine Robertson, Jennifer Goyder and Reanne Cayenne for their feedback throughout the research and writing process of this working paper.

About the Author

Christelle Tessono is a technology policy researcher currently pursuing her graduate studies at the University of Toronto's Faculty of Information. She recently completed a two-year research fellowship as an inaugural Emerging Scholar at Princeton University's Center for Information Technology Policy. She is interested in tackling the relationship between racial inequality and digital technology from a policy lens. As a Digital Policy Hub graduate fellow, she will build on projects she has worked on related to political advertising on social media platforms, gig work, facial recognition technology, data governance and AI regulation.

Works Cited

Attard-Frost, Blair, Ana Brandusescu and Kelly Lyons. 2024. “The Governance of Artificial Intelligence in Canada: Findings and Opportunities from a Review of 84 AI Governance Initiatives.” *Government Information Quarterly* 41 (2): 1–24. www.sciencedirect.com/science/article/pii/S0740624X24000212.

Attard-Frost, Blair and Helen A. Hayes. 2023. “Valuing Value Chains: On Canadian AI Regulation, Co-Governance, and the Scope of AI Value Chains.” SSRN. https://ssrn.com/abstract=4624840.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major and Shmargaret Shmitchell. 2021. “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?” *FAccT ’21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (March): 610–23. https://doi.org/10.1145/3442188.3445922.

Bender, Emily M. and Alexander Koller. 2020. “Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data.” *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*: 5185–98. https://doi.org/10.18653/v1/2020.acl-main.463.

Bol, David. 2020. “14m trees have been cut down in Scotland to make way for wind farms.” *The Herald*, February 29. www.heraldscotland.com/news/18270734.14m-trees-cut-scotland-make-way-wind-farms/.

Brown, Ian. 2023. “Expert explainer: Allocating accountability in AI supply chains.” Ada Lovelace Institute, June 29. www.adalovelaceinstitute.org/resource/ai-supply-chains/.

Castaldo, Joe. 2024. “Canada’s AI infrastructure does not compute.” *The Globe and Mail*, March 21. www.theglobeandmail.com/business/article-canadas-ai-infrastructure-does-not-compute/.

Champagne, François-Philippe. 2023. Correspondence from the Honourable François-Philippe Champagne, Minister of Innovation, Science and Industry, to Joël Lightbound, Chair of the Standing Committee on Industry and Technology – Amendments to AIDA. November 28. www.ourcommons.ca/content/Committee/441/INDU/WebDoc/WD12751351/12751351/MinisterOfInnovationScienceAndIndustry-2023-11-28-Combined-e.pdf.

Dobbs, Graham and Jake Hirsch-Allen. 2024. “Can Canada Compute? Policy Options to Close Canada’s AI Compute Gap.” The Dais, March. https://dais.ca/reports/can-canada-compute/.

Edwards, Dustin, Zane Griffin Talley Cooper and Mél Hogan. 2024. “The making of critical data center studies.” *Convergence: The International Journal of Research into New Media Technologies* 0 (0). https://doi.org/10.1177/13548565231224157.

Erni, Sandy, Xianli Wang, Tom Swystun, Stephen W. Taylor, Marc-André Parisien, François-Nicolas Robinne, Brian Eddy, Jackie Oliver, Brad Armitage and Mike D. Flannigan. 2024. “Mapping wildfire hazard, vulnerability, and risk to Canadian communities.” *International Journal of Disaster Risk Reduction* 101 (February): 104221. https://doi.org/10.1016/j.ijdrr.2023.104221.

Greshake, Kai, Sahar Abdelnabi, Shailesh Mishra, Christoph Endres, Thorsten Holz and Mario Fritz. 2023. “Not what you’ve signed up for: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection.” arXiv. https://doi.org/10.48550/arXiv.2302.12173.

Hadi, Muhammad Usman, Qasem Al Tashi, Rizwan Qureshi, Abbas Shah, Amgad Muneer, Muhammad Irfan, Anas Zafar et al. 2023. “A Survey on Large Language Models: Applications, Challenges, Limitations, and Practical Usage.” TechRxiv. https://doi.org/10.36227/techrxiv.23589741.v1.

Innovation, Science and Economic Development Canada. 2023. “Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems.” September. https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems.

Kahil, Nadine. 2023. “The hidden costs of AI: Thirsty machines in America’s heartland.” *Wired*, September 11. https://wired.me/technology/thirsty-ai-machines-america/.

Laranjeira de Pereira, José Renato. 2024. “The EU AI Act and environmental protection: the case for a missed opportunity.” Heinrich Böll Stiftung, April 8. https://eu.boell.org/en/2024/04/08/eu-ai-act-missed-opportunity.

Leslie, David and Francesca Rossi. 2023. “Generative Artificial Intelligence.” Association for Computing Machinery Tech Brief, Issue 8. https://dl.acm.org/doi/pdf/10.1145/3626110.

Li, Pengfei, Jianyi Yang, Mohammad A. Islam and Shaolei Ren. 2023. “Making AI Less ‘Thirsty’: Uncovering and Addressing the Secret Water Footprint of AI Models.” arXiv. http://arxiv.org/abs/2304.03271.

Lu, Yiwen. 2023. “Generative A.I. Can Add $4.4 Trillion in Value to Global Economy, Study Says.” *The New York Times*, June 14. www.nytimes.com/2023/06/14/technology/generative-ai-global-economy.html.

Luccioni, Alexandra Sasha, Sylvain Viguier and Anne-Laure Ligozat. 2023. “Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model.” *Journal of Machine Learning Research* 24 (253): 1–15. https://www.jmlr.org/papers/volume24/23-0069/23-0069.pdf.

OECD. 2022. “OECD Framework for the Classification of AI Systems.” OECD Digital Economy Papers No. 323. February. https://doi.org/10.1787/cb6d9eca-en.

Office of the Prime Minister. 2024. “Securing Canada’s AI Advantage.” News release, April 7. www.pm.gc.ca/en/news/news-releases/2024/04/07/securing-canadas-ai.

Reese, Justin T., Daniel Danis, J. Harry Caufield, Tudor Groza, Elena Casiraghi, Giorgio Valentini, Christopher J. Mungall and Peter N. Robinson. 2024. “On the limitations of large language models in clinical diagnosis.” medRxiv. https://doi.org/10.1101/2023.07.13.23292613.

Schaeffer, Rylan, Brando Miranda and Sanmi Koyejo. 2023. “Are Emergent Abilities of Large Language Models a Mirage?” *Advances in Neural Information Processing Systems* 36 (December): 55565–81.

Smith, Greg, Michael Bateman, Remy Gillet and Eystein Thanisch. 2023. “Environmental Impact of Large Language Models.” *Amplify* 36 (8): 40–49. www.cutter.com/article/environmental-impact-large-language-models.

Solaiman, Irene, Zeerak Talat, William Agnew, Lama Ahmad, Dylan Baker, Su Lin Blodgett, Hal Daumé III et al. 2023. “Evaluating the Social Impact of Generative AI Systems in Systems and Society.” arXiv. https://doi.org/10.48550/arXiv.2306.05949.

Strome, Elissa. 2024. “Evidence.” House of Commons Standing Committee on Industry and Technology. 1st Sess, 44th Parl, Meeting No. 108. February 5. www.ourcommons.ca/DocumentViewer/en/44-1/INDU/meeting-108/evidence#Int-12552024.

Tessono, Christelle, Yuan Stevens, Momin M. Malik, Sonja Solomun, Supriya Dwivedi and Sam Andrey. 2022. “AI Oversight, Accountability and Protecting Human Rights: Comments on Canada’s Proposed Artificial Intelligence and Data Act.” The Dais, November. https://dais.ca/reports/aida/.

Toronto Metropolitan University, The Dais and the Centre for Media, Technology and Democracy. 2023. “Submission on the Proposed Artificial Intelligence and Data Act.” The Dais, November. https://dais.ca/reports/submission-on-the-proposed-artificial-intelligence-and-data-act/.

Webster, Paul. 2023. “Six ways large language models are changing healthcare.” *Nature Medicine* 29: 2969–71. https://doi.org/10.1038/s41591-023-02700-1.