LEAF FAEJ

WOMEN'S LEGAL EDUCATION & ACTION FUND

FONDS D'ACTION ET D'ÉDUCATION JURIDIQUE POUR LES FEMMES

Submission to The Standing Committee on Industry and Technology on 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*

September 11, 2023

Prepared by:

Rosel Kim
Senior Staff Lawyer, Women's Legal Education
and Action Fund (LEAF)

Dr. Kristen Thomasen
Assistant Professor, UBC

I.   INTRODUCTION

The Women's Legal Education and Action Fund (LEAF) supports the federal government's initiative to regulate the use and impact of artificial intelligence (AI) through the proposed *Artificial Intelligence and Data Act*, which was introduced as part of Bill C-27.[1] However, we are concerned about the absence of a human rights approach within the current proposal, including the lack of recognition of the collective harms that can result from the use of AI, and the range of different institutions and AI-systems that can cause such harms.

While AI has been touted by industry as an innovative tool that will yield benefits for the public, examining the impact of AI from a substantive equality perspective reveals profound harms. As a leading national organization with a mandate to advance substantive gender equality, LEAF urges the government to centre substantive equality and human rights as the guiding principles when regulating the growing use of AI. With this goal in mind, LEAF submits that the scope of AIDA must - at least - be substantially expanded in order to enable regulations that can protect against <u>all</u> present and emerging harms from AI.

Overview of Recommendations:

1. Government institutions must be included in the scope of AIDA (remove s. 3)
2. The statutory definitions of "harm" and "biased output" must be expanded (amend s. 5)
3. Harm mitigation measures must not be restricted to "high-impact" systems (remove s. 7 and remove "high-impact" from ss. 8, 9, 11, 12; amend s. 36(b) so that different obligations for different types of systems can be developed in regulations)
4. "Persons responsible" for AI-systems must explicitly include those involved in system training and testing (amend s. 5)
5. "Persons responsible" should be required to perform an equity and privacy audit to evaluate the possibility and likelihood of harm and biased outputs *in advance* of using, selling, or making available an AI-system. This audit must also be published and made available to the public (amend ss. 8 and 11; amend s. 36 to allow the Governor in Council to outline the requirements for an equity and privacy audit).

---

[1] 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: <https://www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading>.

6. Substantive equality and public consultation must inform the development of regulations (amend preamble and s. 35(1))

About LEAF

LEAF is a national, charitable, non-profit organization that works towards ensuring the law guarantees substantive equality for all women, girls, trans, and non-binary people. LEAF has developed expertise in the gendered and intersectional impact of technology-facilitated violence through intervening in landmark cases before the Supreme Court of Canada[2] and making submissions to Parliament to highlight gender equity implications of online hate.[3] Representatives of LEAF are frequently invited to speak about technology and substantive gender equality, including consultations with representatives of Parliament.

II.   RECOMMENDATIONS TO COMMITTEE

1. Government institutions must be included in the scope of AIDA

> Recommendation: Remove section 3 of AIDA so the Act and subsequent regulations apply to government institutions.

AIDA currently does not apply to government institutions as they are defined in the *Privacy Act*.[4] Further, it does not apply to a product that is under the direction or control of the heads of government agencies responsible for national security such as the Minister of National Defence, the Director of the Canadian Security Intelligence Service.[5] This is problematic from an equality and human rights perspective, given the evidence of government agencies using AI in ways that have discriminated against and/or violated human rights of groups who are already marginalized and over-surveilled. We echo the recommendations made by other

---

[2] LEAF was an intervener in *R v Jarvis*, 2019 SCC 10, where it urged the Supreme Court of Canada to apply an equality lens when interpreting the *Criminal Code* provision on voyeurism.

[3] See e.g. Moira Aikenhead, Suzie Dunn, & Rosel Kim, "Submission to Canadian Heritage on the Federal Government's Proposed Approaches to Address Harmful Content Online" (25 September 2021) online: *Women's Legal Education and Action Fund (LEAF)* <https://www.leaf.ca/submission/leaf-submission-to-canadian-heritage-on-online-hate/>.

[4] *Artificial Intelligence and Data Act*, s 3(1) [*AIDA*].

[5] AIDA, s 3(2).

organizations to the Committee to broaden AIDA's framework to include government institutions.[6]

For example, a 2021 joint investigation by the Office of the Privacy Commissioner and the privacy commissioners of British Columbia, Alberta, and Quebec, found that the Royal Canadian Mounted Police used facial recognition services of Clearview AI, which was found to have breached both federal and provincial privacy laws.[7]

In 2020, two Somali refugees who fled sectarian and gender-based violence were stripped of their refugee status, alleging that the Canada Border Services Agency used Clearview AI to generate photo comparisons between the refugees and two Kenyan citizens.[8] The Minister of Citizenship and Immigration refused to provide details on the investigation, including how CBSA obtained the photo comparisons, citing the *Privacy Act*.[9] The Federal Court found the decision to strip the two applicants of their immigration status to be unreasonable, and noted it was problematic for the Refugee Protection Division (who stripped the applicants' refugee status) to accept the Minister's reliance on the *Privacy Act* to not disclose the details of the investigation.[10] The human rights implications and consequences of government institutions engaging in unlawful and discriminatory AI activity are greater for people facing multiple intersecting systemic barriers.

Given the high rate of error and the serious breaches of the right to safety and equality for marginalized people, the use of AI by government institutions such as law enforcement and border agencies must be included in AIDA so that government agencies are required to assess and take efforts to mitigate risk of harm and biased outputs, and to publicly share information

---

[6] See for e.g. Christelle Tessono et al., "AI Oversight, Accountability and Protecting Human Rights: Comments on Canada's Proposed *Artificial Intelligence and Data Act*" (November 2022) at 4, 9, online: *Cybersecure Policy Exchange* <https://static1.squarespace.com/static/5e9ce713321491043ea045ef/t/63614c030e02403d54fce254/1667320848453/AIDACommentary.pdf>; Professional Institute of the Public Service of Canada (PIPSC), "Brief to the House of Commons Standing Committee on Industry and Technology (INDU) On Bill C-27" (August 2023) at 1-2, online: *House of Commons* <https://www.ourcommons.ca/Content/Committee/441/INDU/Brief/BR12565509/br-external/ProfessionalInstituteOfThePublicServiceOfCanada-e.pdf>.

[7] "Joint Investigation of Clearview AI, Inc. by the Office of the Privacy Commissioner of Canada, the Commission d'accès à l'information du Québec, the Information and Privacy Commissioner for British Columbia, and the Information Privacy Commissioner of Alberta" (2 February 2021), online: *Office of the Privacy Commissioner of Canada* <https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-businesses/2021/pipeda-2021-001/> [*Clearview Investigation*].

[8] *Barre v Canada (Citizenship and Immigration)*, 2022 FC 1078 [*Barre v Canada*].

[9] *Ibid* at para 7.

[10] *Ibid* at paras 79 and 43.

about the steps taken. A 2018 study examining gender classification algorithms from major technology companies found that the algorithms performed the worst (most inaccurately) on dark-skinned women.[11] In their joint investigation of Clearview AI published in 2021, the provincial and the federal privacy commissioners noted how "facial recognition technology has been found to have significantly higher incidences of false positives or misidentifications when assessing the faces of people of colour, and especially women of colour, which could result in discriminatory treatment for those individuals."[12]

As the scope of AIDA is currently limited to interprovincial and international trade and commerce, additional regulation of government uses of AI will also be necessary. But at this stage in the development of a legal framework, it is important, particularly through a substantive equality lens, to see government uses of AI prioritized in legal oversight.[13]

While the tools used by government agencies may improve in accuracy over time, we also caution that simply improving the accuracy of AI without changing the structure of accountability for decisions to use AI will not prevent inequitable outcomes or harm. As LEAF noted in its joint submission with Citizen Lab to the Toronto Police Services Board: "[m]ore accurate and pervasive surveillance or policing supported by AI is not necessarily better for all people, particularly those who are members of already over-policed communities."[14] Having more accurate AI will not prevent state actors from potentially collecting personal information about, and infringing upon the right to privacy and security of people who are already over-surveilled and face risk of discrimination and/or deportation. Accordingly, regardless of the trajectory of the technology, it is crucial that government agencies *at least* come within the scope of the risk-mitigation framework established under AIDA.

---

[11] Joy Buolamwini & Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification" (2018) 81 Proceedings of Machine Learning Research 1-15 <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>. A re-audit in 2019 found that while there was a decrease in error rates in the algorithms than the original audit, it still found that Amazon's facial recognition technology had a 31.4% error rate for darker female faces. See Inioluwa Deborah Raji & Joy Buolamwini, "Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products" (2019) Conference on Artificial Intelligence, Ethics, and Society, online: *Massachusetts Institute of Technology* <https://www.media.mit.edu/publications/actionable-auditing-investigating-the-impact-of-publicly-naming-biased-performance-results-of-commercial-ai-products/>.

[12] *Clearview Investigation* at para 95.

[13] We appreciate as well the guidance provided by the Directive on Automated Decision Making, while emphasizing that further legal scaffolding is needed to ensure AI systems are developed, trained, and used in ways that do not undermine, and ideally support, substantive equality.

[14] Kristen Thomasen & Suzie Dunn,"LEAF and The Citizen Lab Joint Submission on Toronto Police Use of AI Policy" (December 2021) at 9, online: *LEAF* <https://www.leaf.ca/wp-content/uploads/2021/12/TPSB_Public_Consultation_-_AI_Technologies_Policy_-_Dec_15_2021_Submission.pdf>.

2. **The statutory definitions of "harm" and "biased output" must be expanded**

In AIDA, "harm" is defined in terms of the individual body and individual property, and individual economic loss.[15] "Biased output" as it relates to AI is defined as a content, decision, or recommendation that adversely differentiates in relation to an individual based on one or more of the prohibited grounds established in the *Canadian Human Rights Act*.[16] We also note that AIDA takes a risk-mitigation approach to mitigating potential harm and bias (e.g. s. 8). In other words, the legislation is framed to encourage the identification of risks of harm or biased outputs from specific 'high-impact' AI-systems, and delineates steps to mitigate these risks (see our additional recommendations for s. 8 below).

We commend the drafters for understanding individual harm to include both physical and psychological harm, as well as the harms that flow from discriminatory bias. However, the current definitions of harm and biased output do not allow for the recognition of many harms that AI systems can cause to groups, or on the basis of group identity, and do not account for the varying ways in which these harms can arise.[17] These narrow definitions are a missed opportunity for the federal government to recognize and ideally begin to correct existing structural inequities that are reflected in AI, including racism, misogyny, colonialism, and stigma against poverty.

a. Harm to Collective Rights and Interests

<table>
<tr><td>

Recommendation: Amend s. 5 (suggestions in bold)

"Harm means

(a) physical or psychological harm to an individual **or identifiable group**;

(b) damage to an individual's property, **collectively owned property, land or buildings held on behalf of a group or collective, or public property or public spaces**; or

(c) economic loss to an individual **or identifiable group**."

</td></tr>
</table>

---

[15] AIDA, s. 5.
[16] *Ibid*.
[17] Teresa Scassa, "Regulating AI in Canada: A Critical Look at the Proposed *Artificial Intelligence and Data Act*" (2023) 101:1 Can B Rev 1 at 18-19; Kristen Thomasen, "Safety in Artificial Intelligence & Robotics Governance in Canada" (2023) 101:1 Can B Rev 62 at 66.

Collective rights are those held by a group as a whole, in contrast to individual rights which are held individually by members of the group. Collective rights protect the interests of a group, such as cultural and language rights, collective privacy, environmental rights, and labour and union rights, all of which are significantly threatened by the introduction of AI systems.[18] The harms of AI extend beyond the individuals directly impacted by its use, by shifting the norms and expectations regarding the groups to which those individuals belong. As Blair Attard-Frost's submission to the Committee notes, "AI systems are capable of causing a broad range of collective harms, including (but not limited to) collective harms related to human rights violations, collective harms related to IP rights violations, collective harms related to impacts on workers, and collective harms related to environmental impacts."[19] For example, generative AI programs producing artwork for free lead to significant devaluation of work by artists generally.

The current framing of AIDA does not provide for protection against collective or group harm in the forthcoming regulations. This must be rectified through simple amendments to AIDA that would permit the Minister and Governor in Council to consider and protect collective interests in the forthcoming regulations. Additionally, the committee and Minister must widely and meaningfully consult with Indigenous groups to ensure principles and practices of Indigenous data sovereignty are not impeded by, or are exempted from, AIDA while simultaneously determining how the harm to data sovereignty caused by development of AI systems can be explicitly addressed.[20] Furthermore, as emphasized in the Bailey, Burkell, and McPhail submission, Bill C-27's approach to privacy rights and protections must be strengthened in order to mitigate the ways in which AI-systems will erode individual and collective rights, extending beyond privacy right to include equality, expression, and other fundamental rights.

While individual personal harm is sometimes easier to identify than collectively experienced harms, at the time of this submission numerous examples of AI-driven collective harms are

---

[18] See e.g. Jane Bailey, "Towards and Equality Enhancing Conception of Privacy" (2008) 31:2 Dal LJ 267.

[19] Blair Attard-Frost, "Generative AI Systems: Impacts on Artists & Creators and Related Gaps in the Artificial Intelligence and Data Act - Submission to the Standing Committee on Industry and Technology" (5 June 2023) at 13, online: *House of Commons* <https://www.ourcommons.ca/Content/Committee/441/INDU/Brief/BR12541028/br-external/AttardFrostBlair-e.pdf>.

[20] See e.g. Animikii, "#Data Back: Asserting and Supporting Indigenous Data Sovereignty" (2022): online <https://databack.animikii.com/book>; Keoni Mahelona, Gianna Leoni, Suzanne Duncan, Miles Thompson, "OpenAI's Whisper is another case study in Colonisation" (24 January 2023), online: <https://blog.papareo.nz/whisper-is-another-case-study-in-colonisation/>; Maggie Walter et al, "Indigenous Data Sovereignty in the Era of Big Data and Open Data" (2021) 56 Australian Journal of Social Issues 143-156.

taking place. For example, labour movements arising throughout North America in the summer of 2023 highlight the ways in which groups can experience economic and personal harm and discriminatory bias through the adoption of AI-systems. The systems' repeated treatment of workers as replaceable sources of labour that management can control and dismiss undermines workers' humanity and equality as a group. The Standing Committee has already received submissions discussing the harms of AI on artists and creators.[21]

Management decisions to use AI for the sake of efficiency can exacerbate the power imbalance between workers and employers and result in increased precarity for workers. Frontline and service workers have faced increased automated surveillance and threats to their job security - even when participating in legally protected action - when AI has been used to automate and/or manage their work. In July 2023, permanent workers of a hotel in southern California began labour action against the use of automated management tools to replace striking workers. When a worker who signed up for a shift through an app refused to cross the picket line upon realizing he was hired to replace striking workers, he was subsequently penalized on the app by having his account suspended.[22]

The service, frontline, and care workforces are gendered and racialized; ensuring that AI protects the rights of workers is crucial from a gender and racial equity perspective. Failure to consider collective interests will *also* affect the public that seeks out goods and services from such workforces. These harms may be economic, but may also be psychological or physical. Earlier this year, for instance, United States' National Eating Disorder Association (NEDA) made the controversial decision to replace its hotline staff that had recently voted to unionize with a chatbot. This replacement caused economic losses to the unionized employees. It *also* created the possibility for physical and psychological harm for clients accessing the hotline. An activist using the chatbot revealed that "[e]very single thing [the chatbot] suggested were things that led to the development of my eating disorder."[23] This example illustrates how using AI can

[21] Blair Attard-Frost, "Generative AI Systems: Impacts on Artists & Creators and Related Gaps in the Artificial Intelligence and Data Act - Submission to the Standing Committee on Industry and Technology" (5 June 2023) at 13, online: *House of Commons* <https://www.ourcommons.ca/Content/Committee/441/INDU/Brief/BR12541028/br-external/AttardFrostBlair-e.pdf>.

[22] Alex N. Press, "Southern California Hotel Workers Are on Strike Against Automated Management", *Jacobin* (July 2023), online: <https://jacobin.com/2023/07/southern-california-hotel-workers-strike-automated-management-unite-here>.

[23] Chloe Xiang, "Eating Disorder Helpline Disables Chatbot for 'Harmful' Responses After Firing Human Staff", *Vice News* (30 May 2023), online: <https://www.vice.com/en/article/qjvk97/eating-disorder-helpline-disables-chatbot-for-harmful-responses-after-firing-human-staff>.

cause further marginalization of groups who are gendered, rendered vulnerable in different contexts, and likely to be exploited: frontline care workers and people with eating disorders.[24]

Failure to mitigate economic or personal harm to the labour force will allow for the perpetuation of a range of collective harms to workers. Moreover, the harms of AI are likely to more seriously affect workers who are marginalized due to their immigration status, race, and employment status. In other words, where AI shifts institutional and systemic power, it can cause the exact types of harm that AIDA seeks to mitigate but on a collective level - which might currently be seen by those responsible for AI as falling outside of their mitigation obligations. We caution that this could accordingly thwart the government's own objectives within this legislation.

We support recommendations to broaden the jurisdiction of AIDA beyond the trade and commerce power, to further enable the regulatory framework to ensure protection of collective and individual human rights.[25] We noted loss of employment, labour rights violations, and the loss of effective care as collective harms that can arise through the introduction of AI-systems into different contexts. Additional concerns include mass surveillance and the loss of collective privacy, violations of principles of Indigenous data sovereignty through the creation, training, and use of AI-systems, the potential for mass violations of individual and collective human rights and loss of life and liberty as a result of the development of AI tools for military, policing, and judicial systems, and loss of environment caused by the creation and use of of AI-systems. Many of these collective harms are already occurring and have been detailed through research and investigations.[26] Extensive development will be needed in the regulations to ensure that AIDA serves the public interest.

---

[24] On women being over-represented in care work, see Cee Strauss, "Basic Income and the Care Economy" (September 2021) at 21, online: *LEAF* <https://www.leaf.ca/wp-content/uploads/2021/09/Basic-Income-The-Care-Economy-Full-Report-Final.pdf>. On the prevalence of eating disorders in women and girls, see e.g. House of Commons, "Eating Disorders Among Girls and Women in Canada: Report of the Standing Committee on the Status of Women" (November 2014) at 9, 41st Parliament, 2nd Sess, online: *House of Commons* <https://www.ourcommons.ca/Content/Committee/412/FEWO/Reports/RP6772133/feworp04/feworp04-e.pdf>.

[25] Blair Attard-Frost, "Generative AI Systems: Impacts on Artists & Creators and Related Gaps in the Artificial Intelligence and Data Act - Submission to the Standing Committee on Industry and Technology" (5 June 2023) at 13-14, online: *House of Commons* <https://www.ourcommons.ca/Committees/en/INDU/StudyActivity?studyActivityId=12157763>

[26] By way of just some examples, see: Kate Crawford, *Atlas of AI: Power, Politics and the Planetary Costs of Artificial Intelligence* (New Haven: Yale University Press, 2021); Kate Crawford and Vladan Joler, "Anatomy of an AI System", online: <https://anatomyof.ai/index.html>; Emma Strubell, Ananya Ganesh, & Andrew McCallum, "Energy and Policy Considerations for Deep Learning in NLP" (July 2019), 57th Annual Meeting of the Association for Computational Linguistics (ACL), Florence, Italy, online: <https://arxiv.org/abs/1906.02243v1>; Emily Bender et al, "On the Dangers of Stochastic Parrots: Can

For now, we recommend at least that the government must ensure that collective and group interests also come within regulatory scope so that, as the regulations are developed, legislative flexibility is present to ensure that more can be done to address these and other pressing human rights concerns.

b. Broader Understanding of 'Biased Outputs'

<div style="border:1px solid black; padding:10px">

Recommendation: Amend s. 5 (suggestions in bold)

"biased output means content that is generated, or a decision, recommendation, **classification, categorization,** or prediction that is made, by an artificial intelligence system and that adversely differentiates, directly or indirectly and without justification, in relation to an individual on one or more of the prohibited grounds of discrimination set out in section 3 of the Canadian Human Rights Act, or on a combination of such prohibited grounds, **or on the basis of other grounds that result in greater inequality and marginalization for the affected party**. It does not include content, or a decision, recommendation or prediction, the purpose and effect of which are to prevent disadvantages that are likely to be suffered by, or to eliminate or reduce disadvantages that are suffered by, any group of individuals when those disadvantages would be based on or related to the prohibited grounds."

</div>

AI-systems are known to cause not just bodily and property harm, but also discriminatory harm arising from biased outputs. We commend the government for recognizing this in the current draft of AIDA. However, the framing and definition of "biased outputs" in the legislation is too narrow to properly mitigate against the ways AI-systems can lead to unfair discrimination. There are at least two ways in which the current definition is too narrow: it fails to explicitly capture discrimination by proxy; and it overlooks grounds of discrimination (e.g., poverty) that are not covered in human rights legislation.

AI can produce discriminatory outcomes even when it does not directly differentiate based on prohibited grounds. For example, algorithms often classify individuals as risky or safe by analyzing "proxy data" such as where the person lives, other members of their social network,

---

Language Models be Too Big?" FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, at 610–623. And see notes 18, 19, 21-23 *supra*.

and their credit scores.[27] This kind of risk assessment process can label someone as risky by virtue of their socioeconomic status and/or community relationships.[28] The effect of this categorization by proxy is the perpetuation of stigma and discrimination.

For example, many human resources departments have used AI for screening job applicants and making hiring decisions, which has resulted in biased decisions where the AI program learned and replicated gender and racial bias that exist in our society. Amazon had to shut down an AI tool it was developing for making hiring decisions when the company realized the program learned to prefer male candidates over female ones, scoring resumes of graduates from women's colleges lower, and favouring words commonly found in male engineers' resumes over other candidates.[29] Presuming that such a system would be considered 'high-impact' if used, the operator should be obliged to mitigate such biased output. However, the ways in which the system recognized groups of words within a resume, not gender directly, reflects the possibility that proxies would take this instance out of the scope of AIDA, as these words superficially reflect involvement in certain activities or attendance at particular universities. However, as was recognized in the Amazon example, those words serve as a proxy for gender.

Discrimination on the basis of postal code, address, or social networks would seemingly not be captured by AIDA as currently drafted, yet are well-known sites of AI discrimination that reflect recognized grounds of discrimination. These kinds of classifications can cause substantial harm to those who lose out on benefits or are scrutinized and surveilled as a result of an AI-system's recommendation or assessment solely based on their location or communities. AIDA must be sufficiently robust to capture these instances of discrimination by proxy, as AI-systems may frequently render a biased output, not on the explicit basis of a ground enumerated in the *Canadian Human Rights Act*, but on the basis of a proxy for a discriminatory ground.

Moreover, we urge the committee to amend AIDA to recognize discrimination on additional grounds to those set out in the *Human Rights Code*. As noted in our co-authored submission to the Toronto Police Services Board, poverty is another ground of discrimination that will continue to be significant in the context of AI-systems that must be included in the scope of

---

[27] Please note, for this reason, we have also recommended explicitly including "classification" and "categorization" in the definition of biased output to ensure that such methods are captured, as seems to be the legislative intent.

[28] Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Harlow, England: Penguin Books, 2017).

[29] Jeffrey Dastin, "Amazon scraps secret AI recruiting tool that showed bias against women" *Reuters* (10 October 2018), online: <https://www.reuters.com/article/us-amazon-com-jobs-automation-in...-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>.

regulatory oversight.[30] Given the potential for AI-systems to pick up on different, sometimes unpredictable trends in data and to reinforce existing discrimination reflected in those trends, we encourage the committee to broaden the scope for future regulations to consider grounds of discrimination additional to those already recognized in the *Canadian Human Rights Act*.

3. Harm mitigation measures must not be restricted to "high-impact" systems

<div style="border:1px solid black; padding:10px;">

Recommendation: remove s. 7 and remove "high-impact" from ss. 8, 9, 11, 12.

Amend s. 36(b) so that different obligations for different types of systems can be developed in regulations

"Regulations — Governor in Council
36 The Governor in Council may make regulations for the purposes of this Part, including regulations

(b) establishing criteria ~~for the purpose of the definition~~ *high impact system* ~~in subsection 5(1)~~ **for the oversight and management of AI-systems by persons responsible**;

</div>

Currently, AIDA only requires harm mitigation measures for high-impact systems. Classifying systems such that only certain systems receive regulation oversight is a mistake. It incentivizes developers and users of AI to characterize their systems as falling into a less "impactful" and thus unregulated category of AI systems. This is unworkable, as any AI system that could be used in relation to individuals, groups or societies can have the potential to cause harm or a biased output depending on the context for which it was developed and in which it is ultimately used. Many instances of recorded AI bias, including that with serious implications for human rights and substantive equality (examples of which are cited above), have arisen in contexts that the developer did not anticipate. Accordingly, unless "high impact" is defined so broadly as to include almost any system with social valence (which would call into question the need for such a category in the first place), it will not sufficiently address the real concern at the heart of this legislation - mitigating harm and discriminatory bias.

---

[30] See e.g. Kristen Thomasen & Suzie Dunn,"LEAF and The Citizen Lab Joint Submission on Toronto Police Use of AI Policy" (December 2021) at 6-7, 9, online: *LEAF* <https://www.leaf.ca/wp-content/uploads/2021/12/TPSB_Public_Consultation_-_AI_Technologies_Policy_-_Dec_15_2021_Submission.pdf>; see also, e.g. *Virginia Eubanks, Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (New York: St. Martin's Press, 2018)

Instead, all systems should be required to adhere to some degree of oversight and mitigation measures. Simpler systems or those used in very narrow circumstances, or contexts with minimal social impact (e.g. certain industrial settings where a system collects and analyzes data about mechanics/infrastructure), where the likelihood of harm is reasonably anticipated to be low may either more readily adhere to requirements, or could be treated differently in regulations (e.g. less onerous versus more onerous oversight and documentation requirements). Creating a legal structure that only applies to a category of systems considered "high-impact" inappropriately constrains the scope of this law, threatening to render it irrelevant to actual lived experience with AI systems. For the same reasons we noted above with regard to the potential for AI systems to cause harm, we urge the committee to provide greater flexibility for the regulations to set out a range of requirements for operators of different types of systems.

4. "Persons responsible" for AI-systems must explicitly include those involved in system training and testing

Recommendation: Amend s. 5

"Person responsible
(2)   For the purposes of this Part, a person is responsible for an artificial intelligence system, including a high-impact system, if, in the course of international or interprovincial trade and commerce, they design, **train, test,** develop or make available for use the artificial intelligence system or manage its operation."

Algorithms must be trained on data that is "accurate and representative of the subject matter being studied" in order to be effective.[31] If not, the outputs of the algorithm will reflect the flaws and biases of the data it receives - also known as the "garbage in, garbage out" principle.[32] Due to existing inequalities and biases, it is more challenging to ensure that algorithms will receive representative data on issues that are gendered (such as eating disorders) and/or racialized (such as verifying the identities of Black asylum seekers). Accordingly, if the legislative goal of mitigating biased outputs and harm are to be achieved, AIDA must require that persons responsible are accountable for the training and testing methodologies for AI models, in addition to designing and developing them.

---

[31] Kate Robertson et al., "To Surveil and Predict: A Human Rights Analysis of Algorithmic Policing in Canada" (September 2020) at 31, *Citizen Lab*: <https://citizenlab.ca/wp-content/uploads/2020/09/To-Surveil-and-Predict.pdf>
[32] *Ibid*.

Given the high risk of harm, especially for marginalized communities, that can flow from training on non-representative data, or testing that prioritizes accuracy for majority communities without considering the impact on marginalized communities, testing and training methods and data must be explicitly included within the scope of AIDA and regulatory oversight.

5. "Persons responsible" should be required to perform an equity and privacy audit to evaluate the possibility and likelihood of harm and biased output *in advance* of using, selling, or making available an AI-system. This audit must also be published and made available to the public.

<div style="border:1px solid black; padding:10px;">

<u>Recommendation</u>: Amend ss. 8 and 11; Amend s. 36 to allow the Governor in Council to outline the requirements for an equity and privacy audit.

"8 A person who is responsible for an ~~high-impact~~ **AI** system must, in accordance with the regulations, establish measures to identify, assess and mitigate the risks of harm or biased output that could result from the **development or** use of the system. **This must include performing an equity and privacy audit as prescribed by regulation."**

"11 (1) A person who makes available for use a ~~high-impact~~ system must, in the time and manner that may be prescribed by regulation, publish on a publicly available website a plain-language description of the system that includes an explanation of

(a) how the system is intended to be used;

(b) the types of content that it is intended to generate and the decisions, recommendations or predictions that it is intended to make;

(c) the results of an equity and privacy audit as prescribed by regulation; …"

"36 The Governor in Council may make regulations for the purposes of this Part, including regulations

(x) outlining the requirements for equity and privacy audits of AI systems.

</div>

It takes time and effort to detect and challenge the harm of AI. For many people, they may not become aware they were subjected to harm by AI, unless they have capacity to follow up and investigate. Even if they find out they were harmed by AI, pursuing redress from the harm is either challenging or impossible, due to fear of reprisal, or the lack of resources and/or knowledge to pursue legal remedies (where available) Therefore, it is crucial that as many harms as possible from AI systems be mitigated in advance of implementation.

AIDA is structured in a way that seeks to mitigate harm in advance (see our comments on the scope of "harm" above). However, as currently framed, s. 8 does not explicitly require an assessment of risk to take place in advance of the development and use of an AI system. These stages may require separate assessments of risk. Additionally, in conjunction with recommendation 3 above, we urge the government to be explicit in making equity and privacy risks part of the overall mandatory assessment that must occur before the development of a new system. Public and expert consultation must also be a part of the development of requirements for equity and privacy audits.

6. Substantive equality and public consultation must inform the development of regulations

Recommendation: amend preamble and s. 35(1)

Preamble: "And whereas this Act aims to support the Government of Canada's efforts to foster an environment in which Canadians can seize the benefits of the digital and data-driven economy and to establish a regulatory framework that supports and protects Canadian norms and values, including the right to privacy **and substantive equality**;"

"Advisory committee
35 (1) The Minister may establish a committee to provide the Minister with advice on any matters related to this Part. **The committee should reflect a range of perspectives, ensuring public input and consultation particularly from communities most affected by the use of AI systems.**"

AIDA was drafted without substantial public consultation. This approach must be amended in the development of the regulations, which will define the bulk of harm mitigation obligations for the developers and users of AI systems. It is crucial that the government engage in meaningful and in-depth consultation with the public, including civil society organizations and

representatives from communities who are directly impacted by biased outputs of AI, and not just the AI industry or those who stand to benefit from the development of AI systems, in order to adequately consider the types of mitigation requirements and constraints that should be developed in regulations.

Additionally, we echo the recommendation from other groups to establish an independent regulator with appropriate oversight and enforcement powers.[33]

Contact

For further information concerning this submission, please contact:

Kristen Thomasen, Assistant Professor, Faculty of Law, UBC (thomasen@allard.ubc.ca)
Rosel Kim, Senior Staff Lawyer, LEAF (r.kim@leaf.ca)
Pam Hrick, Executive Director and General Counsel, LEAF (pam.h@leaf.ca)

---

[33] Christelle Tessono et al., "AI Oversight, Accountability and Protecting Human Rights: Comments on Canada's Proposed *Artificial Intelligence and Data Act*" (November 2022) at 6-8, online: *Cybersecure Policy Exchange* <https://static1.squarespace.com/static/5e9ce713321491043ea045ef/t/63614c030e02403d54fce254/1667320848453/AIDACommentary.pdf>; OpenMedia, "Submission to INDU Committee Study of Bill C-27, *The Digital Charter Implementation Act, 2022* (3 May 2023) at 20, online: *House of Commons* <https://www.ourcommons.ca/Content/Committee/441/INDU/Brief/BR12448086/br-external/OpenMedia-e.pdf>.