



Hon. Peter Bethlenfalvy
Minister of Finance
President of the Treasury Board
Minister Responsible for Digital and Data Strategy

opendata@ontario.ca

June 4, 2021

Via email

Re: Response to Consultation on [Ontario's Trustworthy Artificial Intelligence \(AI\) Framework](#)

Dear Minister Bethlenfalvy,

Thank you for the opportunity to engage with you and the provincial government in building a digital economy that is powered by trustworthy AI. The Vector Institute and the Schwartz Reisman Institute for Technology and Society (SRI) each have missions to enable the responsible adoption of AI. We are excited by the Trustworthy AI Framework that is being built and are pleased to offer any help we can in developing and executing on this framework.

Our understanding is that the framework would govern the use of AI by government in terms of three important goals: transparency (“no AI in secret”), trustworthiness and fairness (“AI Ontarians can trust”), and the public good (“AI that serves all Ontarians.”) These goals mirror the mission of the Schwartz Reisman Institute to ensure that powerful technologies like AI benefit *all* of humanity and the Vector Institute’s mission to drive research excellence and leadership in AI to foster economic growth and improve the lives of Canadians. We believe in the power of AI and in the critical importance of effective governance to ensure that AI is safe, trustworthy, and fair. We offer here some suggestions for actions to execute on these commitments, and we would welcome the opportunity to provide more in-depth assistance going forward.

At the outset, we want to emphasize that governance for the use of AI both in the private and public sectors is still in early stages across the globe. As documented in a study authored by SRI in December,ⁱ few jurisdictions have yet implemented laws or regulations governing AI. This means that there are few models to follow, but it also means that Ontario has the opportunity to show global leadership. We are enthusiastic about Ontario's willingness to lead on one of the biggest policy issues of the decade.

Given the dynamic nature of AI and the nascent maturity of AI governance, we believe it's important to establish a culture of learning and continuous improvement to develop effective policy for trustworthy AI that protects Ontarians and encourages a thriving economy around AI. We believe that this approach can help achieve transparency, trustworthiness, fairness, and public good from AI while supporting robust investment in AI technologies. In that spirit we offer the following suggestions:

1. **Begin with pilots to identify what to count as "AI."** It is difficult to establish clear definitions for what counts as "AI." A wide variety of computer algorithms are already in use by the Ontario provincial governments and there's no simple test for what counts as AI. New York City's experience with an effort to regulate automated decision systems in city government is a cautionary tale: these efforts largely foundered due to the inability among diverse stakeholders to settle on a definition of "automated decision systems."ⁱⁱ Rather than trying to create an all-purpose definition, we'd suggest piloting the scope of the Trustworthy AI Framework on well-defined high-stakes government decisions. These might include applications in policing, health, and criminal justice.
2. **Emphasize auditing as a regulatory tool.** The novelty of AI systems and the challenge of regulating them make it very difficult to identify the benefits, risks, and the best ways to balance the two up front. For this reason, we believe it is important to focus resources on regular and active auditing of deployed systems, rather than on conducting exhaustive, one-time, up-front reviews. This would be a different approach than that often taken in the domain of data governance, for example, where the focus is on up-front impact assessments, and there is minimal active external auditing once a data governance structure is in place. An auditing approach to regulation will also spur careful attention to developing the metrics and principles that will guide AI policy.
3. **Audit for multiple metrics.** As part of an auditing approach to regulation, we suggest tracking multiple metrics to determine outcomes and identify any unintended consequences from AI deployments in the Ontario government.

Consider an example of applying AI to improve education. In addition to student test scores, it is also important to measure factors of educational success such as engagement, graduation rates, mental health, post-secondary engagement, and employment, as well as to analyze variability and discrepancy by gender, race, language spoken, and other attributes. AI systems are so new that there is still much we don't know about their impact, both positive and negative. What counts as "fair" machine learning, for example, is contested and complex. This field, which has its roots in work done by researchers now affiliated with Vector and Schwartz Reisman,ⁱⁱⁱ has demonstrated several important results about how to assess and mitigate algorithmic discrimination along racial, gender, or other lines. But these results also show that it's not possible to satisfy all definitions of fairness in general (e.g. equality of opportunity and predictive parity)^{iv} and that there are tradeoffs among fairness, quality, and privacy.^v

4. **Focus on building an agile, evidence-based risk framework.** As you are aware, the EU has recently announced proposed legislation to govern AI.^{vi} The EU approach rightfully urges a risk-based approach to regulation with the goal of trying to balance risks and benefits. Although this is a bold first step in a domain where regulation is needed, we see some significant shortcomings in their approach that we would advise Ontario to try to avoid. Of note is the effort to create a definitive list of "high-risk" AI systems and to apply proposed regulations only to systems on that list. The challenge here is that the process of risk assessment has not been based on real-world experiences and concrete metrics. The EU's approach, while not finalized, currently appears to identify entire domains (education, credit, employment, law enforcement, etc.) that certainly involve high-stakes decisions for individuals, but that do not necessarily and pervasively impose high risks of large-scale bad outcomes. We believe that risk assessment for AI systems needs to be grounded in real-world experiences and needs to adapt rapidly as the field advances and AI is applied in new areas. This is also a reason for adopting some of the approaches we advise above, including piloting, auditing, and tracking multiple metrics.

Finally, we believe there's a tremendous opportunity to build on Ontario's Trustworthy AI Framework to spark private sector innovation in governance. We believe that the technical complexity, breadth of applications, and rapid change in techniques related to AI systems makes its governance unlike that of traditional information and software systems. Because of this, conventional regulatory frameworks may be ineffectual or even counterproductive when applied to AI systems. This is a key reason why we strongly suggest taking an approach that emphasizes continuous learning and harnessing private sector innovation.

A key goal, we believe, should be to establish partnerships between the public and private sectors to foster technology investments in safe AI techniques and to build private sector certification regimes that are accountable, reliable, and efficient. We believe there's a timely opportunity to create a legal framework to encourage the emergence of independent third-party certification bodies and to generate the incentives required for entrepreneurs to develop easy-to-implement technologies that verify compliance with government oversight. Schwartz Reisman has recently established a partnership with the University of Toronto's Creative Destruction Lab to foster entrepreneurship in the domain of regulatory technology and, together with the Vector Institute, is working to create open-source tools to improve the safety and security of AI systems. We would welcome the opportunity to help Ontario set the pace globally in developing the legal and economic environment needed to build this important sector for future growth.

Thank you for the opportunity to provide input on Ontario's Trustworthy AI Framework. Vector and SRI share the province's interest in fostering AI that is safe, fair, and trustworthy. We believe that piloting the approach in well-defined domains and focusing on ongoing review and learning will be important foundations for the effort. We believe this will in turn create an opportunity for Ontario to foster an innovative model of private sector certification for AI that promises to become an important new sector. We hope you will call on our expertise as you continue this important work.

Sincerely,



Garth Gibson
President & CEO

Vector Institute
MaRS Centre, West Tower
661 University Ave., Suite 710
Toronto, ON M5G 1M1



Gillian Hadfield
Director & Schwartz Reisman Chair in
Technology and Society

Schwartz Reisman Institute for
Technology and Society, University of
Toronto
MaRS Centre, Heritage Building
101 College St., Suite 230-1
Toronto, ON M5G 1L7

The [Vector Institute](#) drives excellence and leadership in Canada's knowledge, creation, and use of AI to foster economic growth and improve the lives of Canadians. Vector's [Three-Year Strategy](#) aims to advance AI research, increase adoption in industry and health through programs for talent, commercialization, and application, and lead Canada towards the responsible use of AI. [Programs for industry](#), led by top AI practitioners, offer foundations for applications in products and processes, company-specific guidance, training for professionals, and connections to workforce-ready talent. Vector is funded by the Province of Ontario, the Government of Canada through the CIFAR Pan-Canadian AI Strategy, and industry sponsors.

Created in 2019 as a result of a historic gift from Canadian entrepreneurs Gerry Schwartz and Heather Reisman, the [Schwartz Reisman Institute for Technology and Society](#) (SRI) explores and addresses the ethical and societal implications of technology. In its first five years, the Institute is focused on AI, which is rapidly transforming our economic and social environments and in an increasing number of instances threatening or causing harm to human flourishing, social stability, and the well-being of vulnerable populations. At the same time, the potential public benefits of AI are frequently blocked by legal and regulatory obstacles that are poorly adapted to the modern digital environment. SRI was created for the purpose of doing the research and developing the new approaches needed to maximize the benefits and minimize the harms flowing from powerful technologies. SRI's ongoing initiatives include a [series of expert commentaries](#) on the features and implications of privacy law reforms for a digital world, an ongoing collaboration with the Rockefeller Foundation and Stanford University on [innovating AI governance](#), and co-chairing with the Responsible AI Institute a working group on global AI certification as part of the World Economic Forum's Global AI Action Alliance.

ⁱ Zhang, Lawrence. "[Initiatives in AI Governance.](#)" a pre-read paper from the Schwartz Reisman Institute for Technology and Society for the workshop "Innovating AI Governance: Shaping the Agenda for a Responsible Future" (December 2020).

ⁱⁱ Fox Cahn, Albert. "[The first effort to regulate AI was a spectacular failure.](#)" Fast Company (November 26, 2019).

ⁱⁱⁱ Dwork et al. "[Fairness Through Awareness.](#)" Proceedings of the Innovations in Theoretical Computer Science (ITCS) Conference (January 2012).

^{iv} Kleinberg et al. "[Inherent trade-offs in the fair determination of risk scores.](#)" in Proceedings of the Innovations in Theoretical Computer Science (ITCS) Conference (November 2016).

^v Pannekoek and Spigler. "[Investigating Trade-offs in Utility, Fairness and Differential Privacy in Neural Networks](#)" (2021).

^{vi} European Commission. "[Proposal for a Regulation laying down harmonised rules on artificial intelligence](#)" (2021).