JUNE 2021

IPC Comments on the Ontario Government's Consultation on *Ontario's Trustworthy Artificial Intelligence (AI) Framework*

Patricia Kosseim
Commissioner

Information and Privacy
Commissioner of Ontario
Commissaire à l'information et à la
protection de la vie privée de l'Ontario

Over the past two years, the Government of Ontario (the "Province") has sought feedback on its emerging approach to artificial intelligence. This has included its 2019 *Consultation for Ontario's Digital and Data Strategy*,[1] its 2020 *Artificial Intelligence (AI) Guidance* "Alpha Principles" for Ethical Use and Transparency,[2] and most recently, its consultation on *Ontario's Trustworthy Artificial Intelligence (AI) Framework* ("Framework")*.*[3]

In its recently published *Digital and Data Strategy*,[4] the Province expresses its intention "to build a digital economy powered by ethical artificial intelligence (AI) rooted in democratic principles and individual rights." This strategy is referenced in the consultation web page for the Framework, with the Framework being presented as a foundational element to support this strategy's realization. The proposed Framework is centered on three high-level draft commitments, with each of these commitments supported by three potential actions. The draft commitments are as follows:

1. **No AI in secret:** The use of AI by the government will always be transparent, with people knowing when, why, and how algorithms are used and what their rights are if harm occurs.
2. **AI use Ontarians can trust:** Rules and tools are in place to safely and securely apply algorithms to government programs and services based on risk.
3. **AI that serves all Ontarians:** Ontarians benefit economically and socially from AI technologies that are rooted in individual rights and reflect the diverse communities across the province.

These are important commitments for the Province to make, and are generally aligned with the Information and Privacy Commissioner of Ontario's (IPC) mandate of protecting privacy and promoting transparency. In line with our strategic priority *Privacy and Transparency in a Modern Government,* our goal is to advance Ontarians' privacy and access rights by working with public institutions to develop bedrock principles and comprehensive governance frameworks for the responsible and accountable deployment of digital technologies.[5] As part of an open and ongoing dialogue around the use of AI in Ontario, we hereby offer our initial thoughts about the use of AI in the public sector as a response to the Province's consultation on the Framework.

[1] Government of Ontario. *Consultation for Ontario's Digital and Data Strategy*. Webpage last updated January 2021. **https://www.ontario.ca/document/consultation-ontarios-digital-and-data-strategy**

[2] Government of Ontario. *Artificial Intelligence (AI) Guidance*. March 30, 2021. **https://www.ontario.ca/page/artificial-intelligence-ai-guidance**

[3] Government of Ontario. *Consultation: Ontario's Trustworthy Artificial Intelligence (AI) Framework*. May 5, 2021. **https://www.ontario.ca/page/ontarios-trustworthy-artificial-intelligence-ai-framework-consultations**

[4] Government of Ontario. *Building a Digital Ontario*. April 30, 2021. **https://www.ontario.ca/page/building-digital-ontario**

[5] Information and Privacy Commissioner of Ontario. *IPC Strategic Priorities 2021-2025*. April 22, 2021. **https://www.ipc.on.ca/about-us/ipc-strategic-priorities-2021-2025/**

Our submission includes several general considerations with respect to the overall framing and scope of the proposed Framework. We then turn to each commitment, and offer considerations we feel will strengthen the Framework's treatment of access and privacy issues. We are strongly committed to working closely with the Province to ensure the rights of privacy and access are core components of a provincial AI governance model.

GENERAL CONSIDERATIONS

The IPC is supportive of the broad commitments proposed in the Framework, which should help hold the government to account for its use of AI. We also recognize that further elaboration and specificity is required in order for the Framework to accomplish its intended outcomes.

Given the high-level nature of the Framework's commitments, our comments that follow draw from our own experiences with, and research on, AI and related technologies. We also identify areas for further consideration as the government continues its efforts on this topic. We intend to keep our considerations general, rather than setting out specific policy directions or proposals. We look forward to a regular dialogue with the Province on the topic of the Framework.

1. CLEARLY DEFINE KEY CONCEPTS AROUND AI AND WHICH CONCEPTS ARE SUBJECT TO THE FRAMEWORK

Developing a governance model for AI must include as a foundational element a clear definition of AI and related concepts. In the absence of precise definitions, ambiguity and misunderstanding may emerge that lead to gaps in accountability and risks going unidentified. A definition must also be flexible enough to accommodate future technological developments. Therefore, the Framework should include clear definitions of key concepts relating to AI.

Following is a list of definitions of some of the key components or concepts of AI. These definitions are intended to help distinguish some terms that are commonly used in an interchangeable manner. We do this for the principal purpose of clarifying our comments that follow but would further welcome the opportunity to contribute to any initiative of the Province to formalize any related definitions. Key AI related terms include the following:

- **Artificial Intelligence** in use today involves the use of computation to analyze certain types of data according to a generalized model of the world to accomplish defined objectives by generating outputs that impact the external environment.[6] AI in use today and in the near future is not what is known in the scientific literature as "artificial general intelligence" – where machines can think as humans do and exhibit self-awareness.[7] Contemporary AI systems are typically narrowly focused on a clear problem domain and adopt a particular approach to accomplishing their objectives. Components of AI implementations include:

  - o An **AI model** is a set of instructions on how to interpret a particular subject that can either be explicitly crafted by humans (classical AI) or developed through machine learning. **Machine learning** uses statistical or other numerical approaches to build a model based on **training data** in a way that does not requiring explicit programming by a human. Models can be used for purposes including to make **predictions**, to put data into categories through **classification**, and to **generate** original data that bears similarities to real world examples.
  - o An **AI system** is an implementation of one or more AI models into a computer system that is implemented in an environment in order to accomplish a particular objective. The computer system may include a role for human operators, or operate relatively autonomously.
  - o The **environment** in which an AI system is deployed is a key consideration. The environment may be observed by an AI system (i.e. data input), and affected by actions the system takes (i.e. output). AI systems may perform differently in different environments.

- An **algorithm** is a set of instructions developed to solve a particular problem. Algorithms can be implemented in computer code but do not have to be. They are used in mathematics, in spreadsheets, and in all aspects of computer software, AI included. It is important to differentiate algorithms from AI since algorithms are so ubiquitous. AI models are instances of (often very complex) algorithms. As an example of this distinction, the evaluation of an application for a government benefit against a pre-defined set of criteria would likely be considered an algorithm, but not an AI system.

- **Automated Decision Making (ADM)** is defined in the Government of Canada's *Directive on Automated Decision-Making* as "any technology that either assists or replaces the judgement of human decision-makers," and employs techniques such as those leveraged in AI models (including statistical and linguistic

[6] This definition is based on the model definition of AI presented in the European Commission. *Proposal for a Regulation laying down harmonised rules on artificial intelligence.* Recital 6. April 21, 2021. **https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence**

[7] See, e.g. Santus, E., Christin, N. and Jayarm, H. "Artificial Intelligence." *Technology Factsheet Series.* Belfer Center for Science and International Affairs. Harvard Kennedy School. January 2020. **https://www.belfercenter.org/publication/technology-factsheet-artificial-intelligence**

methods).[8] Some definitions include consideration that the ADM system impacts "opportunities, access, liberties, rights, and/or safety."[9] ADM systems are important to delineate from AI since AI can be employed in contexts outside of administrative decision making, where its use may have less of an impact on individuals or groups.

The Framework should clarify whether it applies to AI systems in general, or only systems that meet a narrower definition of automated decision making. For instance, would AI systems that do not analyze information about, make decisions about, or otherwise impact individuals be in scope?

Similarly, the line between statistical analysis (a longstanding government practice) and AI is not necessarily a clear one. Absent an established scope (and given the use of both 'algorithm' and 'AI' in the Framework), it is unclear whether the Framework is intended to apply to any data-centric government processes, or only AI systems.

We do not suggest that there is a "right" approach to the scope of technologies covered by the Framework – justifications exist for both broad application (e.g. individuals' rights with respect to a decision should be largely independent of what specific technologies were used to make it) and narrow application (e.g. governance frameworks can be more specific and directed when applied to a narrower set of processes). However, whatever choices are made about the proper scope of the Framework, it is important that all stakeholders share a common understanding.[10]

2. CLEARLY DEFINE WHICH INSTITUTIONS WILL BE SUBJECT TO THE FRAMEWORK

AI, machine learning, and automated decision making are in widespread use today. The technologies are being tested or deployed by a broad range of institutions and for a variety of purposes across the spectrum of provincial public sector programs, including:

- **Health care:** St. Michael's Hospital in Toronto has implemented a machine learning model that analyzes a wide range of data including historical emergency department visits, weather patterns, and scheduled events in the area to predict how many patients will visit the hospital's emergency department on a given day.[11]

[8] Government of Canada. *Directive on Automated Decision-Making*. Last modified April 1, 2021. **https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592**

[9] Richardson, R. "Defining and Demystifying Automated Decision Systems." *Maryland Law Review* (pre-print). March 26, 2021. **https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3811708**

[10] Note: we have referred exclusively to AI and AI systems in this submission; this is done for the sake of clarity, as opposed to recommending a particular scope for this framework

[11] Unity Health Toronto. Strategic Plan 2019-2024. April 20, 2019. **http://bce.unityhealth.to/unity-health-toronto-strategic-plan-2019-2024.pdf**; Invest Ontario. *Spotlight: Toronto hospital prescribes AI to cure ER wait times*. February 10, 2020. **https://www.investontario.ca/spotlights/toronto-hospital-prescribes-ai-cure-er-wait-times**

- **Policing:** Law enforcement agencies in Ontario are exploring the use of facial recognition technologies, which are powered by computer vision, a branch of AI that interprets the visual world and identifies and categorizes objects based on images.[12] The IPC is actively engaged on the issue of police use of facial recognition as part of our strategic priority, which focuses on *Next-Generation Law Enforcement*.

- **Education:** Some remote university exams are being monitored by software that uses machine learning to detect and predict potential academic misconduct. For instance, these systems can use a variety of AI models that categorize objects and activities based on data obtained through real-time monitoring of a variety of sources, such as a student's mouse and keyboard activity, video footage from their webcams, and audio from their computer microphones.[13]

- **Transportation:** The Ontario government partnered with organizations to use AI to detect the number of people in vehicles driving in high occupancy toll lanes on provincial roads.[14]

- **Digital service delivery:** Business owners can interact with the "Grants Ontario Chatbot" to obtain information about funding opportunities from the provincial government.[15] Chatbots often use natural language processing, a type of AI which is designed to interpret and generate language and in some cases generate a realistic-seeming dialogue with humans.

Clearly, AI can be employed in numerous ways, for different purposes, and with different intended users of the technology.

Given the widespread use of AI across various sectors, the Framework should clarify the program areas and/or ministries that will be subject to these commitments. For instance, much of the language in the consultation paper suggests a focus on AI systems that are public-facing. However, consideration should also be given to AI systems that may be used for other back-end tasks including policy development, planning and forecasting, and cybersecurity. Such activities often take place outside of the public view, and in some cases, justifications may exist for a certain degree of confidentiality (as is the case with certain cybersecurity measures.)

[12] Chellappa, R., Sinha, P. and Phillips, P.J. "Face Recognition by Computers and Humans." *IEEE Computer*. February 2009. **https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=903088**

[13] Graham, A. "As concerns linger, Western University promises solution to remote exam proctoring software." *Global News*. March 14, 2021. **https://globalnews.ca/news/7693767/western-university-proctortrack-concerns/**

[14] Government of Ontario. *News Release: Ontario Enhancing Government Services Through Partnerships with Small Business.* November 9, 2017. **https://news.ontario.ca/en/release/46968/ontario-enhancing-government-services-through-partnerships-with-small-business**

[15] Government of Ontario. *Get funding from the Ontario government.* February 14, 2020. **https://www.ontario.ca/page/get-funding-ontario-government#section-4**

3. ENSURE THE FRAMEWORK APPLIES THROUGHOUT THE ENTIRE AI LIFECYCLE

When put to use, AI, like other technologies, can be understood as progressing through a ‘lifecycle’ of discrete stages: from initial concept, to design, implementation, transitioning to ongoing maintenance and finally decommissioning. Each stage in the lifecycle is marked with particular problems to address, actions to take, and issues for which due care must be applied.

Institutions may benefit by approaching trustworthy AI from a lifecycle perspective. Such an approach can help to ensure that risks to access, privacy, and public trust are identified and addressed at appropriate times. A lifecycle model for AI systems put forward by the Organisation for Economic Co-operation and Development[16] includes the following stages:

1) **Design, data, and modeling**: System objectives, underlying assumptions, context, and requirements are specified. Data to power the AI system is then collected, processed, and checked for quality. The AI system developers then create or select a model or algorithm that is trained or calibrated against the data set.

2) **Verification and validation:** Developers assess their model for its performance against objectives. This could include assessing false positives, false negatives, and/or performance under a variety of conditions.

3) **Deployment:** The model and its overall system is launched for use in an environment. The system may begin to monitor the environment, assess collected data using its models, and generate outputs such as predictions, categorizations, decisions, and assessments.

4) **Operation and monitoring:** The AI system is in operation, with its outputs being used in service of the AI system’s objectives. The system is monitored in light of performance and quality evaluation criteria. Based on monitoring results, the system operators may take their system back to earlier phases to re-evaluate the design and training of the system.

Institutions may become involved in AI systems at any point in their lifecycle. AI systems may be built in-house, a vendor product may be customized to government specifications, or a program area may subscribe to a vendor-managed cloud-based AI tool. Regardless of when an institution becomes involved, a consideration of each phase of the lifecycle is important, as risks associated with AI are present during each stage.

[16] Organisation for Economic Co-Operation and Development. “The Technical Landscape.” *Artificial Intelligence in Society*. June 11, 2019. **https://www.oecd-ilibrary.org/sites/8b303b6f-en/index.html?itemId=/content/component/8b303b6f-en**

For these reasons, we suggest that the government clarify that its Framework applies to all stages of the AI systems lifecycle.

COMMITMENT 1: NO AI IN SECRET

The use of AI by the government is always transparent, fair, and equitable.

The IPC strongly supports the notion that government discloses its uses of AI. Such disclosures are aligned with transparency obligations under Ontario's *Freedom of Information and Protection of Privacy Act* (*FIPPA)* and its municipal equivalent, the *Municipal Freedom of Information and Protection of Privacy Act (MFIPPA)*.

FIPPA and *MFIPPA* are premised, in part, on the principle that information about government activity is essential to the proper functioning of open and transparent democratic institutions. As noted by the Supreme Court of Canada, government accountability is supported by access to information legislation that can help the public understand the activities of government.[17]

This commitment is important because AI systems can challenge the ability for the public to get access to information about government decisions and operations in several ways. For instance, machine learning models can often be so complex that they function as "black boxes," where the data used and assessed, and the reasoning behind an automated decision, is not readily understood or documented.[18]

This transparency challenge can be compounded if institutions rely upon models created by third party organizations that withhold details about their models due to intellectual property concerns,[19] or if models are embedded deep within systems and are not properly documented.[20] The difficulty in explaining how AI systems work can lead to difficulties in challenging an institution's compliance with its legal and other obligations, which is of particular consequence in light of widespread concerns around bias and fairness in AI systems.

[17] Supreme Court of Canada. *Dagg v. Canada*. June 6, 1997. **https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/1525/index.do**

[18] Thomas, N., Chochia, E., and Linsday S. "Regulating AI: Critical Issues and Choices." *Law Commission of Ontario.* April 2021. **https://www.lco-cdo.org/wp-content/uploads/2021/04/LCO-Regulating-AI-Critical-Issues-and-Choices-Toronto-April-2021-1.pdf**

[19] Rubenstein, D. "Federal Procurement of Artificial Intelligence: Perils and Possibilities." *The Great Democracy Initiative*. **https://greatdemocracyinitiative.org/wp-content/uploads/2020/12/Artificial-Intelligence-Report_121320-FINAL.pdf** p. 32

[20] See, e.g. Office of the Privacy Commissioner of Canada. *Joint investigation of the Cadillac Fairview Corporation Limited by the Privacy Commissioner of Canada, the Information and Privacy Commissioner of Alberta, and the Information and Privacy Commissioner for British Columbia.* October 28, 2020. **https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-businesses/2020/pipeda-2020-004/**

Transparency of AI systems is key to public trust for a variety of reasons including because concerns respecting the discriminatory impacts of AI systems on marginalized communities are well documented and long standing. Biases embedded in machine learning models are associated with factors including the data used to train them, the design choices of their creators, and the criteria used to evaluate and test their effectiveness. The disparity between the context in which the training data was collected and the environment in which the AI system is deployed can lead to inaccurate inferences and prejudicial decisions being made about individuals and communities.[21]

Under *FIPPA*, institutions have an obligation to take reasonable steps to ensure that personal information in their custody or control is not used unless it is accurate and up-to-date.[22] Individuals also have the right to request correction of their personal information that they deem is inaccurate, and to require that a statement of disagreement be attached to personal information that is not corrected.[23] These responsibilities and rights are challenged when administrative decisions supported by AI systems are not easily understood. It is difficult for an individual to correct biased or discriminatory inferences that are made within the confines of an inscrutable machine learning model. Accuracy and correctness are key elements of the right to privacy that are directly challenged in the context of AI systems.

4. EXPAND THE JUSTIFICATION FOR TRANSPARENCY TO SUPPORT CHALLENGING NOT ONLY BIAS, BUT ALL FORMS OF INACCURACY AS WELL AS THE OVERALL APPROPRIATENESS OF AN AI SYSTEM

Two of the potential actions under the commitment to *No AI in Secret* focus on being transparent when AI is used to make decisions about people, and for people to be able to challenge such decisions if they were made in a biased manner.

We note that regardless of the source of an error (bias or otherwise) any inaccuracy should be contestable. While errors in an AI system may be the result of bias, they may also be the result of inaccuracy in the system's implementation, inaccuracy in the data used to make a decision, or an inappropriate approach to the problem. Similarly, the

[21] See, e.g. Buolamwini, J. and Gebru, T. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." *Proceedings of Machine Learning Research 81:1-15*. 2018. **http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf**

[22] *Freedom of Information and Protection of Privacy Act* s. 40 (2). **https://www.ontario.ca/laws/statute/90f31#BK62.** An institutions that is a health information custodian has a similar obligation under the *Personal Health Information Protection Act, 2004* (*PHIPA*) to take reasonable steps to ensure that any personal health information it uses is as accurate, complete, and up-to-date as is necessary for the purposes for which it uses the information. *PHIPA* s. 11 (1). **https://www.ontario.ca/laws/statute/04p03#BK16**.

[23] *Freedom of Information and Protection of Privacy Act* s. 47 (2). **https://www.ontario.ca/laws/statute/90f31#BK72.** Individuals have a similar right under *PHIPA* to request correction of their personal health information that they deem is inaccurate, and to require that a statement of disagreement be attached to personal health information that is not corrected. *PHIPA* s. 55. https://www.ontario.ca/laws/statute/04p03#BK77.

right to explainability is not intended solely to protect against bias, but is an overall element of procedural fairness.

Beyond ensuring an avenue to contest individual decisions, transparency in the use of AI systems also creates the opportunity for individuals to challenge the appropriateness of specific uses of AI. In some instances, the use of an AI-powered system may in-and-of itself have significant impacts on populations or groups of people (for instance, by enabling greater surveillance), even if challenges associated with accuracy and bias in specific instances can be overcome.

We therefore ask the government to consider expanding its justification for transparency to support challenges beyond countering bias to a broader range of reasons why the public might seek to understand and, as needed, contest the use of and outcomes generated by AI systems.

5. EXPAND THE SCOPE OF ITS TRANSPARENCY COMMITMENTS

In our view, the third potential action under the first commitment (providing clarity and transparency to the public on how Ontario collects data for use in algorithms) should be expanded to ensure Ontarians have the opportunity to understand not only *that* data is being collected, but *what* data is collected.

For example, individuals should be provided an opportunity take steps to review and ensure the accuracy of any data collected for use as part of a decision making process. As discussed above, accuracy is a privacy principle that the IPC believes will need to take on a heightened role in the age of AI and data analytics.

To assess the appropriateness of a given AI system, individuals must also know the purpose for which it is being used and whether it achieves the intended purpose. To address this, the Framework should include a requirement to make public the purposes for each AI system and develop mechanisms to publicly demonstrate the effectiveness of a system in furtherance of its objectives.

When evidence indicates that the AI system is no longer effective in furthering its objectives, individuals should be able to challenge the ongoing use of the system, and call for its cessation.

To conclude our remarks with respect to this commitment, we reiterate our support for an overall move toward openness for government use of AI, but note that transparency should be understood broadly and encompass the entire AI lifecycle. This includes the reasons that AI is adopted, how it is developed, what data it uses (both in development and for specific applications), for which purposes, how it makes decisions or arrives at outcomes, how those decisions are acted upon, and how effective it is in achieving its objectives, etc. Knowledge of the existence of an AI system is an important step, but it is only a first step.

6. INCLUDE A CLEARER FOCUS ON ACCOUNTABILITY

In October 2020, the IPC, along with several other Canadian and international privacy and data protection regulators, sponsored a resolution adopted by the Global Privacy Assembly (GPA) focusing on *Accountability in the Development and Use of Artificial Intelligence*.[24] This resolution notes that accountability is a critical component of the legal and ethical development of AI, and takes the view that accountability obligations should be assessed against clearly defined principles and frameworks.

Transparency is a key component of an overall accountability framework for government activities, but accountability encompasses much more than transparency. Inspired by the international resolution mentioned above, accountability means demonstrable compliance with applicable laws, policies and frameworks, "in particular through the adoption and implementation of appropriate, practicable, systematic and effective measures."[25]

So, while the Framework's first commitment refers tangentially to the concept of accountability, we would urge the government to place greater emphasis on this key principle by elevating it to a self-standing commitment on its own. In support of such a commitment, several potential actions would be needed to establish the building blocks of an effective accountability program, including:

- **Developing a clear internal governance structure for AI**. An internal governance structure would have to provide clear roles and responsibilities and document critical management decisions relating to the use of AI in compliance with the Framework, such as sign-off on risk assessments, approval for human-in-the-loop decision points, and approval of policies and procedures in support of the Framework and its commitments.
- **Appointing a designated role responsible for AI oversight**. This role would oversee an institution's adherence to the Framework, help develop resources and procedures, act as an internal advocate in support of the Framework, and be reachable by the public with questions about the institution's AI practices.
- **Establish standards for engagement and consultation.** Those deploying AI should also understand the limits of their own expertise, and establish criteria for when other parties should be engaged to aid in determining what measures

[24] Global Privacy Assembly. *Adopted Resolution on Accountability in the Development and use of Artificial Intelligence*. 42nd Closed Session of the Global Privacy Assembly. October 2020. **https://globalprivacyassembly.org/wp-content/uploads/2020/11/GPA-Resolution-on-Accountability-in-the-Development-and-Use-of-AI-EN.pdf**

[25] Ibid. See also similar definitions in: Centre for Information Policy Leadership. *What Good and Effective Data Privacy Accountability Looks Like*. May 2020. **https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_accountability_mapping_report__27_may_2020__v2.0.pdf**, and Information Accountability Foundation. *The Essential Elements of Accountability*. January 2019. **https://informationaccountability.org/publications/**

accountability will require. For instance, during the design stage, the Province should consider the suggestion of the Centre for Information Policy Leadership that systems designers consult with internal or external review boards for guidance on high risk systems.[26]

- **Implement whistleblowing/reporting mechanisms**. Such mechanisms would provide a clear channel through which people can report instances of legal non-compliance, unauthorized high risk uses of AI, or failure to adhere to the Framework without fear of reprisal.

We therefore recommend that the government call for appropriate management controls to oversee compliance with the Framework through an additional commitment dedicated to ensuring accountability with respect to the responsible use of AI, along with the related potential actions needed for its effective implementation.

COMMITMENT 2: AI USE ONTARIANS CAN TRUST

Risk-based rules are in place to guide the safe, equitable, and secure use of AI by government.

Privacy and data protection authorities, including the IPC, recognize that AI poses fundamental challenges to numerous principles upon which privacy legislation is based.[27] For instance, the principle of *limiting collection* is challenged by AI, and machine learning in particular, since AI models typically perform best when trained on a large and diverse volume of data. It is also not uncommon for organizations to repurpose already-collected data for use in AI training, challenging the principle of *purpose limitation*.[28] *Limiting retention* is an enigma for machine learning, as training information and insights derived from it may persist in a model long after the underlying training data is deleted or otherwise rendered out of date.[29]

[26] Centre for Information Policy Leadership. *CIPL Recommendations on Adopting a Risk-Based Approach to Regulating Artificial Intelligence in the EU*. March 22, 2021. **https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_risk-based_approach_to_regulating_ai__22_march_2021_.pdf**

[27] International Conference of Data Protection and Privacy Commissioners. "Declaration on Ethics and Data Protection in Artificial Intelligence." *40th International Conference of Data Protection and Privacy Commissioners*. October 23, 2018. **https://globalprivacyassembly.org/wp-content/uploads/2018/10/20180922_ICDPPC-40th_AI-Declaration_ADOPTED.pdf**; Resolution of the Federal, Provincial and Territorial Information and Privacy Commissioners. *Effective privacy and access to information legislation in a data driven society*. October 1-2, 2019. **https://www.priv.gc.ca/en/about-the-opc/what-we-do/provincial-and-territorial-collaboration/joint-resolutions-with-provinces-and-territories/res_191001/**

[28] Centre for Information Policy Leadership. "First Report: Artificial Intelligence and Data Protection in Tension. Artificial Intelligence and Data Protection." *Delivering Sustainable AI Accountability in Practice*. October 10, 2018. **https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_ai_first_report_-_artificial_intelligence_and_data_protection_in_te....pdf**

[29] Izzo, Z. et al. "Approximate Data Deletion from Machine Learning Models." *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS) 2021*. Pre-print. **https://arxiv.org/abs/2002.10077**

In addition, one of the foundational building blocks for privacy protective data sharing – de-identification – is challenged by research showing that machine learning models can successfully re-identify large portions of data sets thought to be de-identified.[30]

These challenges to existing privacy best practices underscore the importance of due diligence with respect to AI systems that may involve personal information. The IPC supports the government's proposed commitment to take a risk-based approach to the use of AI, and the considerations we put forward are intended to help specify how that commitment can be strengthened. We also support tests that examine the robustness, reliability, accuracy, and security of AI systems, including identifying and addressing bias in the systems.

Per the Global Privacy Assembly resolution on *Accountability in the Development and Use of Artificial Intelligence* we sponsored in 2020, the IPC committed to working with organizations to ensure that risks to privacy and access rights, and other human rights, are assessed before AI systems are put into use.[31] Our focus on this work is also in line with our strategic priority of Privacy and Transparency in a Modern Government.

7. DEFINE A CLEAR SCOPE, CRITERIA, AND METHODOLOGY FOR RISK ASSESSMENT AND PUBLISH ASSESSMENT RESULTS

With respect to the potential action to "[a]ssess whether to use an algorithmic assessment tool as a way to measure risk, security, and quality," we observe that numerous forms of assessment may be required to address different categories of risk. Many existing algorithmic impact assessment tools are principally focused on automated decision making systems, and primarily focus on the explainability, auditability, and fairness of the system.[32] These are critical issues for automated decision systems to be assessed against, and we encourage the government to leverage existing algorithmic impact assessment methodologies. However, we would like to ensure that existing privacy risk assessment tools are also leveraged when appropriate.

We observe that the Government of Canada's Algorithmic Impact Assessment (AIA) tool[33] notes that Privacy Impact Assessments may need to be conducted in addition

[30] Rocher, L., Hendrickx, J.M. & de Montjoye Y. "Estimating the success of re-identifications in incomplete datasets using generative models." *Nature Communications* 10:3069. 2019. **https://www.nature.com/articles/s41467-019-10933-3**

[31] See sections 1(1) and 1(2) of the Global Privacy Assembly. *Adopted Resolution on Accountability in the Development and use of Artificial Intelligence*. 42nd Closed Session of the Global Privacy Assembly. October 2020. **https://globalprivacyassembly.org/wp-content/uploads/2020/11/GPA-Resolution-on-Accountability-in-the-Development-and-Use-of-AI-EN.pdf**

[32] Reisman, D. et al. Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability. AI Now Institute, New York University. 2018. **https://ainowinstitute.org/aiareport2018.pdf**

[33] Government of Canada. *Algorithmic Impact Assessment*. Last modified March 22, 2021. **https://open.canada.ca/aia-eia-js/**

to an AIA if personal information is used. Similarly, the algorithmic impact assessment does not assess cybersecurity to a level of assurance that would be realized through established threat/risk assessment and penetration testing methodologies.

We also refer back to our earlier considerations 1 and 3, and note that clear definitions and a lifecycle approach will be critical in developing a process for identifying what types of systems will need to be assessed for risk, and at what point in the lifecycle those assessments should occur. Clear definitions are also important in assigning risk levels in a consistent manner.

The Framework should clarify if a risk-based approach is to be applied for all uses of AI, irrespective of whether they involve the processing of personal information. This clarification should recognize the risk that the use of AI systems may re-identify information previously thought to be de-identified.

In alignment with the commitment of *No AI in secret*, as well as the government of Canada's *Directive on Automated Decision Making*,[34] we also would remind the Province of the importance of keeping records of the risk assessments it conducts and encourage the open publication of risk assessment reports wherever feasible, or at least summaries thereof to the extent appropriate.

8. DEVELOP MECHANISMS TO ENSURE THAT THE PURPOSE OF AN AI SYSTEM DOES NOT SIGNIFICANTLY CHANGE WITHOUT RE–ASSESSMENT

Most risk assessments are 'point-in-time' analyses, meaning that a system's objectives, design, accompanying policies and procedures, and other information are examined for risks and a report issued with recommendations on how to mitigate the risks. Risk assessments often become out-of-date or obsolete as a system changes or if new uses are introduced.

For instance, a facial recognition technology system used by law enforcement in high security facilities might be assessed for risk, and mitigating controls put in place appropriate for that context. However, if law enforcement were to expand the use of facial recognition technology systems beyond high security facilities to allow for more general surveillance across new classes of facilities, risks could increase or new risks could be introduced.

For this reason, we support the Province's proposal to continuously test for bias and risk, and suggest it consider building in mechanisms to trigger the need for re-assessments if certain criteria change, such as when the actual use varies from the original intended use, or if the effectiveness of the system is diminished (in line with our consideration 4).

[34] Government of Canada. *Directive on Automated Decision-Making*. Last modified April 1, 2021. Appendix C. **https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592**

9. ENSURE CRITERIA, FUNDING, TRAINING, INSTITUTIONAL STRUCTURES AND OTHER NECESSARY SUPPORTS ARE IN PLACE FOR REQUIRED HUMAN OVERSIGHT AND INTERVENTION

Human oversight can play an important role in upholding trust in AI. We strongly support a risk-based approach to determining when human oversight and intervention is required in an AI system. We further note that such oversight and intervention plays a role throughout the AI systems lifecycle.

During the regular operations of an automated decision system, the government of Canada's *Directive on Automated Decision Making* requires that higher risk systems cannot make decisions without having specific human intervention points throughout the process, and the final decision must be made by a human. In support of this, the directive also requires that employees be trained so that they are able to "review, explain and oversee" the operations of an automated decision system.[35]

The Province should implement clear criteria for when human oversight and intervention in systems are required, and ensure appropriate funding, training, institutional structures, and other supports are in place to ensure the effectiveness of that oversight and intervention.

10. CLARIFY ALIGNMENT BETWEEN AI STRATEGY AND RELATED LEGISLATIVE FRAMEWORKS AND PROPOSED REFORMS

There is a recognized need to update Canada's privacy laws to address both the barriers to innovation and loopholes in protection that have emerged as a result of the technological developments of the past two decades.[36]

With respect to recent initiatives that affect the legislative oversight of AI, we note that recent amendments to *FIPPA* have created a framework for the sharing of personal information between ministries (and with extra-ministerial data integration units) for certain designated purposes. The extent to which AI is intended to be used in data integration units is not clear at this time.

The Province has also been exploring private sector privacy legislation. A harmonized approach to protecting privacy rights (including access to one's personal information and explanation of decisions impacting them) in the context of automated decision-making in the public and private sectors would help bring coherence and consistency across the AI systems lifecycle, which may include commercial entities at various points.

[35] Government of Canada. *Directive on Automated Decision-Making*. Last modified April 1, 2021. Section 6.3.5. **https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592**

[36] See, e.g. Scassa, T. "Data Protection Law." *Artificial Intelligence and the Law in Canada (eds. Martin-Bariteau, F. and Scassa, T.).* LexisNexis 2021.

The government should clarify how its AI strategy aligns with existing and new laws. We would welcome the opportunity to provide advice on how to achieve this regulatory alignment and consistency.

COMMITMENT 3: AI THAT SERVES ALL ONTARIANS

Government use of AI reflects and protects the rights and values of Ontarians.

As we have discussed, the IPC recognizes that AI systems may affect human rights in a wide variety of ways. We have focused in particular on issues of access to information and the protection of privacy, in keeping with our mandate. However, as discussed above, there are issues relating to the use of AI, and in particular bias, discrimination, and fairness, that extend beyond the IPC's mandate and into other areas of human rights. Ethical concerns that arise when governments replace human decision-making with artificial agents require careful consideration. The Province would benefit by consulting with the Ontario Human Rights Commission, Ombudsman of Ontario, and ethics scholars to address these challenges.

With this in mind, the considerations we put forward under the commitment of *AI that serves all Ontarian*s seek to ensure that the Province 1) implement clear criteria with respect to AI systems posing unacceptable risks, 2) carefully confer and coordinate oversight responsibilities among existing bodies where appropriate, 3) recognize the applicability of many already-existing guidance and governance approaches to aspects of AI systems, and 4) carefully consider and consult on what it truly means to serve all Ontarians.

11. CONSIDER EXPANDING THE CRITERIA BY WHICH CERTAIN USE CASES OF AI ARE PROHIBITED, AT LEAST TEMPORARILY

Numerous initiatives are now underway to prohibit or significantly constrain the use of AI in certain contexts. For instance, the European Union's recently proposed regulation creates a category for 'high-risk' AI systems, which are subject to stricter requirements than other systems. The regulation also strictly prohibits AI practices that run contrary to EU values (e.g. by violating human rights) and thus create an 'unacceptable risk'.[37] As another example, in 2019 the state of California prohibited the use of facial recognition in body-worn cameras used by police services[38] for a three-year period,

[37] European Commission. *Proposal for a Regulation laying down harmonised rules on artificial intelligence.* April 21, 2021. **https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence**

[38] Electronic Privacy Information Center. *State Facial Recognition Policy*. **https://epic.org/state-policy/facialrecognition/**

reportedly to provide time for additional safeguards to be developed around the technology's use.[39]

We are thus glad to see the potential action item of assessing whether the government should prohibit the use of AI in certain use cases where vulnerable populations are at an extremely high risk. However, we would encourage the Province to consider whether there may also other situations in which the use of AI is inappropriate, such as where the human rights of Ontarians in general may be significantly negatively impacted.

We would also direct you to the recent *Ethics, Transparency and Accountability Framework for Automated Decision-Making*[40] guidance document issued by the UK Government's Office for Artificial Intelligence. That document includes the following consideration:

> Before using this framework, you should consider whether using an automated or algorithmic system is appropriate in your context.
>
> Scrutiny should be applied to all automated and algorithmic decision-making. They should not be the go-to solution to resolve the most complex and difficult issues because of the high-risk associated with them.

The Province should include a statement of this nature in its Framework. A commitment to caution with respect to high risk applications of AI, particularly in the absence of strong legislative requirements, oversight mechanisms, or technical controls, would help ensure trust in the Province's AI strategy.

12. CAREFULLY DESIGN INDEPENDENT OVERSIGHT MECHANISMS

Strong and independent oversight should be a key component of the developed Framework. This should include both the ability to perform proactive audit/review of the operations of AI systems,[41] and provide an avenue of redress for individuals wishing to challenge outcomes of AI systems. The burden should not be entirely placed on affected individuals to challenge both individual outcomes of AI systems and to identify broader systemic biases within those systems. In their report *Regulating AI: Critical Issues and Choices*, the Law Commission of Ontario (LCO) expressed support for the

[39] Thebault, R. "California could become the largest state to ban facial recognition in body cameras." *The Washington Post*. September 11, 2019. **https://www.washingtonpost.com/technology/2019/09/12/california-could-become-largest-state-ban-facial-recognition-body-cameras/**

[40] Government of the United Kingdom. *Guidance: Ethics, Transparency and Accountability Framework for Automated Decision-Making.* May 13, 2021. **https://www.gov.uk/government/publications/ethics-transparency-and-accountability-framework-for-automated-decision-making/ethics-transparency-and-accountability-framework-for-automated-decision-making**

[41] Office of the Privacy Commissioner of Canada. *A Regulatory Framework for AI: Recommendations for PIPEDA Reform*. November 2020. **https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/completed-consultations/consultation-ai/reg-fw_202011/**

principle of independent oversight of government AI and automated decision-making systems, but noted it was less certain about the institutional design or placement of that oversight function.[42]

The IPC would welcome an opportunity to collaborate with the government and other stakeholders (such as the LCO) to develop an independent oversight model that is appropriate for purpose. The oversight model should be developed carefully in light of already-existing oversight bodies such as the IPC, the Ontario Human Rights Commission, the Ombudsman, as well as the newly proposed Data Authority.[43] Such coordination will be critical to ensure that roles and responsibilities are as clear, streamlined, and coherent as possible, and to avoid needless redundancy, delay and confusion for individuals seeking to contest inaccurate, unfair, or unreasonable decisions.

13. CONSIDER EXISTING GUIDANCE AND GOVERNANCE FRAMEWORKS

We emphasize that many of the challenges with respect to trustworthy AI are new variations of challenges that government has already addressed. We therefore recommend that the Framework reference and build upon existing regulations, policies, standards, and guidelines where appropriate, both to ensure consistency of approach as well as enabling the government to better focus its efforts on the unique challenges of AI.

For instance, as mentioned above, the principle of accuracy (and hence, non-bias) already figures prominently as an obligation of public institutions covered by FIPPA. Similarly, while explainability is particularly challenging to achieve in some AI systems, the notion of "giving reasons" is already a long-standing hallmark of administrative law. Explainability can also be supported through strong documentation and recordkeeping practices, which government institutions are already expected to have. Transparency of government processes is also not a new concept. Alignment with existing standards, where appropriate, could help demystify AI and normalize it as a component of government operations, subject to an institution's full range of governance mechanisms.

14. CLARIFY WHAT IT MEANS FOR AI TO SERVE ALL ONTARIANS

We would also like to offer one final reflection from our recent process to determine our Strategic Priorities. In our initial consultation paper, we put forward a potential priority around "Responsible Use of Data for Good." In developing this priority, it was

[42] Thomas, N., Chochia, E., and Linsday S. "Regulating AI: Critical Issues and Choices." *Law Commission of Ontario.* April 2021. **https://www.lco-cdo.org/wp-content/uploads/2021/04/LCO-Regulating-AI-Critical-Issues-and-Choices-Toronto-April-2021-1.pdf**

[43] Government of Ontario. *Building a Digital Ontario*. April 30, 2021. **https://www.ontario.ca/page/building-digital-ontario**

made clear to us by our *ad hoc* Advisory Committee[44] and other stakeholders that the concept of "good" was not necessarily clearly defined, and questions needed to be asked such as: What is good? Data for whose good? Who gets to make the ultimate determinations, and who is accountable? And finally, are there boundaries that cannot be crossed, regardless of the good that might be achieved?

This type of inquiry seems consistent with the commitment to *AI that serves all Ontarians* and the associated potential actions. Accordingly, the Province should consider a potential action that addresses these critical questions head on. In engaging with "sector leaders and civil society to develop a standard for 'trustworthy AI' and a process to certify that vendors are meeting the government's standard", the Province should consult more broadly on what it truly means for AI to serve all Ontarians. By undertaking this work, it will create a strong foundation for the other actions – creating clarity for developers about what they should consider as potential harms, or for what a "trustworthy AI standard" is intended to achieve.

CONCLUSION

In conclusion, we commend the government for undertaking this important work and agree that creating guidelines for the government's use of AI represents an important early step towards the overall goal of building a digital economy that is powered by trustworthy AI.

We look forward to engaging with the Province as it advances its work on the proposed Framework and supporting commitments, and we offer to support this work as part of our own strategic priority of championing Privacy and Transparency in a Modern Government.

[44] See Appendix A, Information and Privacy Commissioner of Ontario. *IPC Strategic Priorities 2021-2025*. April 22, 2021. **https://www.ipc.on.ca/about-us/ipc-strategic-priorities-2021-2025/**

Information and Privacy
Commissioner of Ontario

Commissaire à l'information et à la
protection de la vie privée de l'Ontario

2 Bloor Street East, Suite 1400
Toronto, Ontario
Canada M4W 1A8
Telephone: 416–326–3333

www.ipc.on.ca
info@ipc.on.ca

June 2021