Consultation: Ontario’s Trustworthy Artificial Intelligence (AI) Framework

Submission by Huawei Technologies Canada Inc.
June 4, 2021

Introduction

Artificial intelligence (AI) has the potential to unlock untold economic and societal benefits for the people of Ontario. As an innovation leader the province is well positioned to capitalize on this new generation of technology. However, AI, without a proper legal and regulatory regime to support it, leads to concerns around accountability, safety, rights and trust.

As a global leader in ICT technology, including AI, which operates in over 170 countries with vastly different legal, regulatory and cultural environments, Huawei Technologies Inc. (Huawei) is perhaps well positioned to offer some of our observations from our global practices combined with our understanding of the current landscape here in Ontario.

Huawei describes AI in a simplified way as a set of sciences, theories and techniques whose purpose is to create computers and software that are capable of intelligent behavior. This technology development and the governance of it can also be used to reflect ethics and values.

Although in some areas this might be more challenging to achieve, in regards to central provincial policy areas like health care, where AI can help to improve patient care, or education, the technology can immensely help improve quality of life, and ultimately government efficiency and decision making, in ways now being examined by policy makers and leaders.

Below we have provided some information and comments as it relates to the three commitments outlined by the Province. However, above all else, we would encourage the Ontario Government to reach out beyond its border in helping to develop this framework. The reason for this is since the digital economy driven by AI typically involves an international supply chain, a fragmented governance framework may lead to regulatory arbitrage and vicious competition across different regions. With this in mind, establishing multilateral AI governance mechanisms consisting of members from governments, civil society and private-sector is critical to promote a basic consensus of trusted AI across the world and avoid fragmentation of responsibilities globally.

Commitment 1: No AI in secret

When examining potential applications of AI, very few uses pose a “high” risk to the people of Ontario with most being fairly benign in this regard. Furthermore, for some areas of AI a clear binary classification of risks is challenging, particularly as technology develops. Transparency, fairness and equity will be critical to ingrain in the overall framework in order to maintain the public trust needed to more specifically regulate the detailed intricacies involved in some high risk areas such as finance or justice. These principles can lead to more efficient use of AI regulation focused on higher vs. lower risk areas. It will be necessary for the Government and stakeholders to distinguish between these high or low risk scenarios, and when possible avoid those viewed as overly burdensome and entirely unnecessary in order to promote and not hinder innovation.

Some of the other challenges the Government will face here are the intricacies of what is meant by full transparency. Does this go as far as open algorithmic source code, or open training data or is it just merely disclosure of the use of algorithm with a general description? A good principle to operate under is that users have the right to be thoroughly informed with the relevant information when they are using AI applications or interacting with AI systems. When the decisions made by AI systems have significant impacts, it must be ensured that users are aware of the accuracy and limitations of AI decision making, provided with the background information about how the AI system works and how the data is used, and, where appropriate, explanations should be given.

Accountability, will be fundamental as AI further expands in to the day to day aspects of our lives. Some of the concepts discussed in the documentation surrounding this framework mention rights to explainability, contest and others as part of creating this accountability, however these should be considered perhaps as more of a remedy to create procedural fairness against AI bias in order to ensure the non-discrimination principles. Accountability for organizations and individuals is discussed further in the document below.

Commitment 2: AI use Ontarians can trust

Ontario is not alone in addressing the concerns around AI applications in our lives. The reality is that existing international, regional and/or national binding and/or non-binding legal instruments are not currently sufficient to regulate AI systems in order to ensure one can trust AI technology.

Unfortunately, these existing instruments; lack specific principles for the design, development and application of AI systems; do not provide enough guidance to the designers, developers and deployers of AI systems; and; do not provide for rights for those citizens interacting with AI.

When determining whether to adopt a risk based approach, the next logical question should be kept in mind regarding, if this was the path forward, how are these risk levels classified and regulated differently? Difficult legal and ethical questions will face Ontario, however some guiding principles that will ensure trust in AI systems are that:

- Individuals should always have the right that any decision taken by an AI system in the framework of judicial proceedings are reviewed by a “human” judge.
- Individuals should in almost all circumstances have a right to demand the review of an algorithmic based decision by a human being.

Public institutions and corporations will ultimately be on the “frontline” of the interaction between AI and the public, and thus in almost all these organizations there should be an official responsible person for reviewing algorithmic based decisions.

In terms of necessary legislation and regulation, Huawei would suggest the most critical areas that will require the utmost formal and detailed focus by policy makers will be in high risk sectors such as justice, law enforcement, and the ever challenging field of social networking and other internet intermediaries. AI systems within elections and other areas related to the democratic process should also be strictly regulated as loss of public trust in this area could prove cataclysmic.

In regards to some of the lower risk areas, commonly strong and clear guidelines, and voluntary certification programs, can achieve what is needed to make the citizens “trust” an AI system in a certain instance or application. These guidelines that can help regulate the development and deployment of AI should focus on areas like respect for human rights and dignity, non-discrimination, privacy and data protection, legal clarity and certainty, and ultimately an accessible, affordable and transparent system to appeal decisions made by AI systems, particularly in the public realm, with a clear remedy path for errors.

Examples of critical principles that could form part of an AI guideline for are:

- Implementing technical and organizational measures and procedures – proportional to the type of system that is developed – to ensure that data subjects' privacy and personal data are respected, both when determining the means of the processing and at the moment of data processing.
- Assessing and documenting the expected impacts on individuals and society at the beginning of an artificial intelligence project and for relevant developments during its entire life cycle.
- Identifying specific requirements for ethical and fair use of the systems and for respecting human as part of the development and operations of any artificial intelligence system.

When examining the use of algorithmic assessment tools to measure risk, security and quality, and when ensuring processes are in place to test and evaluate algorithms for bias/risk with necessary human oversights, it will be important to outline how this applies to specific scenarios and approach it as such. The rationale for this is these as a general concept or requirement may not exist, or be too broad to judge its implementation in certain more specific scenarios, or in some scenarios there is a likelihood that this principle or concept is simply not applicable.

Ultimately due to the complexity of AI and the extremely broad breadth of its applications, both legal and regulatory changes, as well as voluntary guidelines and certifications, will make up the framework needed for protecting individual rights and ensuring Ontarians trust AI.

Commitment 3: AI that serves all Ontarians

The Government use of AI must reflect and protect the rights and values of Ontarians and ultimately serve them all equally. AI can present risks in forms such as perpetuating discrimination or privacy concerns, and it will be up to the Government to identify where this risk is too high, particularly towards vulnerable populations.

Ultimately, Huawei believes that the use of AI should ensure fairness and justice, avoid prejudice and discrimination against specific groups or individuals, and be sure to not further disadvantage vulnerable populations.

Areas like emotional analysis to measure engagement, deep fakes or cheap fakes, and AI applications aimed at predicting recidivism are examples of where AI systems are perhaps not providing desired outcomes for society. However, the economic and social benefits can be immense including AI applications providing support to the healthcare system via triage or treatment delivery coupled with applications for quicker and more accurate diagnosis, or assisting in the allocation of educational resources, or in examining gender equality, or even in predicting possible consequences and their likelihood from natural disasters and climate change.

Overall, in order to allow benefits and avoid negative outcomes, development and application of AI must be diverse and inclusive, as it must ensure specific individuals or minority groups are not subject to unfair bias, stigmatization, or discrimination.

AI practitioners should strive to minimize the introduction of bias when developing and deploying AI. Such harms can be mitigated through both technical tools and organizational changes; for example through de-biasing, compliance with diversity and discrimination legislation, and training of employees. It is important to note that there is no panacea and there are no one-off technical fixes and the level of risk should be considered when determining applicable desirable methods

AI must not be deployed in ways that will compound the disadvantages of already vulnerable populations. In order to achieve that, AI practitioners should use algorithms and data models that eliminate bias, use training datasets that meet diversity requirements and perform extensive validation of AI systems.

This is not without challenges as they remain in defining areas like “data set requirements”, as who and how will you set the measures of data for representativeness, accuracy, consistency and validity while also avoiding the introduction of bias? It will likely depend on who is asking for these data sets and for what applications. Furthermore, if data is collected with “user privacy” in mind, one runs the risk of counterintuitively reducing the quality of the data set.

As the complexity of Government interaction with its citizens is immense, this relationship, and the implications of AI within the interactions in this symbiotic relationship, will ultimately lay the framework for how the people of Ontario see AI and how they see it adjusting and working to meet their needs and desires. With this in mind, it is critical that there are even higher transparency standards for public entities using AI versus non private organizations. The Government should establish public oversight mechanisms including certification and quality labelling and regular audits and evaluations.

Beyond the needs for more specific regulation and guidelines, ultimately the overarching mandate of these oversight mechanisms should be to protect all Ontarians’ values and rights by:

- Considering individuals’ reasonable expectations by ensuring that the use of artificial intelligence systems remains consistent with their original purposes, and that the data are not used in a way that is incompatible with the original purpose of their collection,
- Taking into consideration not only the impact that the use of artificial intelligence may have on the individual, but also the collective impact on groups and on society at large,
- Ensuring that artificial intelligence systems are developed in a way that facilitates human development and does not obstruct or endanger it, thus recognizing the need for delineation and boundaries on certain uses.

Overall, AI deployment and application must meet the requirements of lawfulness, fairness, security and privacy protection. Logics (such as prediction, judgment, and automated actions) generated by the system must be explainable and transparent. Personal information and data are protected and managed in compliance with the GDPR and other applicable laws. Data and information are shared and AI development is accelerated while privacy is ensured. Unfairness is inherent in society or caused by procedures. To mitigate unfairness inherent in society, AI deployment must ensure fair distribution of benefits and costs, and avoid bias and discrimination.