LCO Final Report

Accountable AI

June 2022

LAW COMMISSION OF ONTARIO
COMMISSION DU DROIT DE L'ONTARIO

About the LCO

The Law Commission of Ontario (LCO) is Ontario's leading law reform agency. The LCO provides independent, balanced and authoritative advice on complex and important legal policy issues. Through this work, the LCO promotes access to justice, evidence-based law reform and public debate.

The LCO evaluates laws impartially, transparently and broadly. The LCO's analysis is informed by legal analysis; multi-disciplinary research; contemporary social, demographic and economic conditions; and the impact of technology.

LCO reports are a practical and principled long-term resource for policymakers, stakeholders, academics and the general public. LCO's reports have led to legislative amendments and changes in policy and practice. They are also frequently cited in judicial decisions, academic articles, government reports and the media.

This report is part of the LCO's ongoing AI, ADM and the Justice System project. The project brings together policymakers, legal professionals, technologists, NGOs and community members to discuss the impact of AI and algorithms on access to justice, human rights and due process. The LCO's current AI initiatives include *Accountable AI, AI and Human Rights* (with the Ontario and Canada Human Rights Commissions) and *AI in the Criminal Justice System*. Earlier LCO AI-related projects are listed on the next page.

The LCO is also undertaking projects addressing the Indigenous Last Stages of Life, consumer protection, protection orders, and environmental accountability.

The LCO is located at Osgoode Hall Law School, York University, Toronto.

More information about the LCO is available at ***www.lco-cdo.org***.

Law Commission of Ontario Reports

Indigenous Legal Issues in the Last Stages of Life (Forthcoming 2022)
Comparing European and Canadian AI Regulation (November 2021)
Legal Issues in the Last Stages of Life (October 2021)
Regulating AI: Critical Issues and Choices (April 2021)
The Rise and Fall of AI and Algorithms in American Criminal Justice: Lessons for Canada (October 2020)
Defamation Law in the Internet Age (March 2020)
Class Actions Objectives, Experiences and Reforms (July 2019)
Legal Capacity, Decision-making, and Guardianship (March 2017)
Simplified Procedures for Small Estates (August 2015)
Capacity and Legal Representation for the Federal RDSP (June 2014)
Review of the Forestry Workers Lien for Wages Act (September 2013)
Increasing Access to Family Justice (February 2013)
Vulnerable Workers and Precarious Work (December 2012)
A Framework for the Law as It Affects Persons with Disabilities (September 2012)
A Framework for Teaching about Violence Against Women (August 2012)
A Framework for the Law as It Affects Older Adults (April 2012)
Modernization of the Provincial Offences Act (August 2011)
Joint and Several Liability Under the Ontario Business Corporations Act (February 2011)
Division of Pensions Upon Marriage Breakdown (December 2008)
Fees for Cashing Government Cheques (November 2008)

LCO AI, ADM and the Justice System Project: Notable Reports and Activities

Comparing European and Canadian AI Regulation (November 2021)

AI Case Study: Probabilistic Genotyping DNA Tools in Canadian Criminal Courts (June 2021)

Regulating AI: Critical Issues and Choices (April 2021)

Legal Issues and Government AI Development (March 2021)

LCO/Ontario Digital Service Workshop (November/December 2020)

The Rise and Fall of Algorithms in the American Justice System: Lessons for Canada (October 2020)

LCO Forum on AI and ADM in the Civil and Administrative Justice System (December 2019)

LCO Forum on AI in Ontario's Criminal Justice System with The Citizen Lab, Criminal Lawyers Association and the International Human Rights Program, Faculty of Law, University of Toronto (March 2019)

AI, Automated Decision-Making: Impact on Access to Justice and Legal Aid (June 2019)

AI for Lawyers: A Primer on Artificial Intelligence in Ontario's Justice System with Element AI and Osgoode Hall Law School (May 2019)

Roundtable on Digital Rights and Digital Society with the Mozilla Foundation (March 2018)

Contacting the LCO

The LCO believes that successful law reform depends on broad and accessible consultations with individuals, communities and organizations across Ontario. As a result, the LCO is seeking comments and advice on this report. As such, the LCO welcomes comments and advice on this report.

There are many ways to get involved. The LCO can be contacted at:

Law Commission of Ontario
Osgoode Hall Law School, York University
2032 Ignat Kaneff Building
4700 Keele Street
Toronto, Ontario, Canada
M3J 1P3

Email: LawCommission@lco-cdo.org
Web: www.lco-cdo.org
Twitter: @LCO_CDO
Tel: (416) 650-8406
Toll-free: 1 (866) 950-8406

Authors and Contributors

Susie Lindsay, Counsel & Principal Author
Nye Thomas, Executive Director
Erin Chochla, Research Lawyer
Andrew Seo, Executive Officer
Noor Malik, Special Projects Assistant
Nadim Dabbous, LCO Student Scholar
Dana Levin, LCO Student Scholar
Damiana Pavone, LCO Student Scholar
Michael Piaseczny, LCO Student Scholar
Holly Reid, LCO Student Scholar
Rob Wright, Graphic Designer

Funders

The Law Commission of Ontario is funded by the ***Law Foundation of Ontario,*** the ***Law Society of Ontario***, ***Osgoode Hall Law School,*** and ***York University***. The LCO also receives significant in-kind support from Osgoode Hall Law School and York University.

The Law Foundation of Ontario
Advancing access to justice

Law Society of Ontario | Barreau de l'Ontario

OSGOODE HALL LAW SCHOOL | YORK U

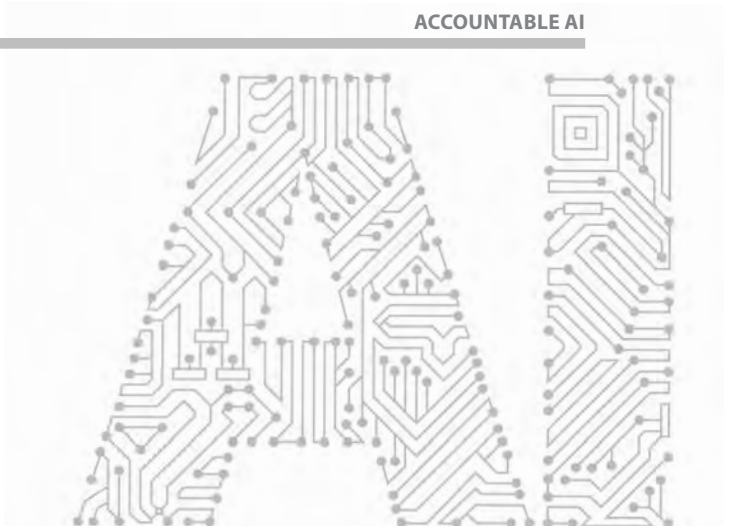
Advisory Panel

Abdi Aidid
Assistant Professor, University of Toronto, Faculty of Law
Raj Anand
Chair of the LCO Board, Partner at WeirFoulds LLP
Amy Bihari
Senior Manager, Government of Ontario
Insiya Essajee
Counsel, Ontario Human Rights Commission
Michelle Mann
General Counsel, Department of Justice
Jonathon Penney
Associate Professor of Law, Osgoode Hall
Carole Piovesan
Partner and Co-Founder, INQ Data law
Marcus Pratt
Policy and Strategic Research, Legal Aid Ontario
Jennifer Raso
Assistant Professor of Law, University of Alberta
Teresa Scassa
Professor of Law, University of Ottawa
Julia Stoyanovich
Assistant Professor of Computer Science and Engineering, New York University
Christiaan van Veen
Director, Digital Welfare State and Human Rights Project, NYU School of Law

Citation
Law Commission of Ontario, *Accountable AI,* (Toronto: June 2022).

CONTENTS

Accountable AI



INTRODUCTION

This is the latest in a series of Law Commission of Ontario (LCO) Issue Papers considering the use of artificial intelligence (AI), automated decision-making (ADM) and algorithms in the Canadian justice system.

This paper considers how to ensure legal accountability when governments and public agencies use AI to make or assist decision-making in the civil and administrative justice systems. More specifically, this paper considers the following issues:

- Why and how are governments using AI to assist decision-making;
- How AI changes government decision-making;
- The relationship between Trustworthy AI and legal accountability;
- The relationship between AI regulation, litigation and access to justice;
- The strategies or reforms needed to promote meaningful legal accountability for government AI decision-making.

The context for this analysis is the extraordinary growth in the use of AI and ADM by governments across the world. The breadth and pace of government AI systems reflects the perceived potential of AI to improve the accuracy, speed and consistency of government decision-making.

Some of the notable uses of AI in government to date include:

- Natural language processing tools to improve the quality of adjudicative decision-writing;
- Child risk assessment tools that predict the potential for child neglect and abuse;
- Immigration detention tools that evaluate whether to detain a person;
- Biometric surveillance systems;
- Teacher evaluation systems that measure a teacher's impact on educational achievement;
- Public benefit fraud detection algorithms;
- Predictive policing systems that predict the potential location of crimes or potential offenders;
- Bail and sentencing algorithms that predict the potential for recidivism;
- Tax compliance algorithms.

Notwithstanding AI's potential, government use of AI is controversial. There are many examples of government AI systems that have proven to be biased, illegal, secretive or ineffective. As a result, many governments – including the Government of Ontario – are adopting "Trustworthy AI" frameworks to assure the public and stakeholders that government AI development and use will be transparent, legal and beneficial.

Trustworthy AI is an important initiative, and the LCO commends the Government of Ontario for publicly committing to this goal. Achieving "Trustworthy AI", however, will depend on governments addressing a complex series of policy, legal and operational questions that go far beyond public statements of principle. Addressed thoughtfully, the answers to these questions will

help governments and public agencies maximize AI's benefits and minimize its harms.

This paper considers a singularly important dimension of "Trustworthy AI": legal accountability for government AI systems.

The LCO believes that widespread adoption of AI systems has the potential to transform the legal and policy landscape of government decision-making. As a result, government AI accountability strategies must be designed to respond to the unique features of AI decision-making.

The LCO's analysis is strongly influenced by "technological due process" principles and priorities. This concept, based on a seminal 2008 article by Professor Danielle Keats Citron, suggests that AI and algorithms require deeper analysis of due process and regulatory issues than traditional legal models may suggest.[1] This concept is grounded in a belief that

> *…the accountability mechanisms and legal standards that govern decision processes have not kept pace with technology. The tools currently available to policymakers, legislators, and courts were developed primarily to oversee human decisionmakers…our current frameworks are not well- adapted for situations in which a potentially incorrect, unjustified, or unfair outcome emerges from a computer.*[2]

This paper addresses the key areas of human rights, administrative law, privacy law and civil procedure to determine if there are gaps or unanswered questions that must be addressed to ensure appropriate legal accountability for government AI systems.[3]

The LCO has concluded that "accountable AI" depends on a mix of law reform tools and strategies, including front end regulation, substantive law reform, enhanced due process protections, and innovative initiatives to improve access to justice. The LCO has also concluded that many tools and strategies are available to policymakers today. Others will depend on policymakers and stakeholders coming together to address a complex series of legal accountability challenges that often combine legal and technical analysis, a combination that itself raises new questions and difficulties. A complete list of the LCO's 19 recommendations are included in Appendix A.

Canadian governments have an opportunity to become leaders in successful AI deployment by applying hard-learned lessons and taking proactive measures to ensure trustworthy and accountable AI. Absent these measures, government ministries, agencies, tribunals and courts will likely need to address important legal and technical issues on a case-by-case basis, resulting in poorer public services, biased and inconsistent government decision-making, diminished rights protection, delays, and unnecessary costs and litigation.

This paper is the latest in a series of LCO Issue Papers addressing AI and ADM in the Canadian justice system. Earlier papers address the use of AI and ADM in the criminal justice system,[4] legal issues and government AI development,[5] how to regulate AI,[6] and the use of AI to generate evidence in criminal proceedings.[7] Throughout this work, the LCO is particularly concerned with access to justice issues and the impact of AI and ADM on vulnerable Ontarians, including racialized communities, low-income Ontarians and Indigenous peoples. Experience demonstrates that AI and ADM systems often have disproportionate and negative effects on these communities.

The LCO's AI-related work reveals consistent themes and law reform proposals across sectors and legal policy areas. This paper applies many of these lessons and themes to the use of AI tools in the civil and administrative justice systems.

Readers should note that this paper directs recommendations to the Government of Ontario. However, the LCO believes our analysis and recommendations are equally applicable to other Canadian governments and a wide range of public agencies and institutions.

SUMMARY

Creating accountable AI is a multi-faceted process. This section summarizes this report's findings, themes and recommendations to support this goal.

Trustworthy AI

All Ontarians have a fundamental interest in ensuring government AI systems are effective, accurate, fair, legal and accountable.

The Government of Ontario has taken many positive steps to promote trust in provincial AI systems, including its emerging *Trustworthy AI Framework* and draft Alpha and Beta documents. More work is needed, however, to "operationalize" these important commitments.

The LCO recommends the provincial government fulfil its *Trustworthy AI* framework by:

- Committing to not deploy high-risk AI systems prior to adopting its *Trustworthy AI Framework;*
- Establishing the *Trustworthy AI Framework* in legislation and regulations;
- Committing to transparency, accountability and public engagement in provincial AI systems;
- Ensuring that criminal justice AI systems (such as facial recognition, biometric identification, predictive policing, and bail/sentencing risk assessments) are included in a dedicated criminal justice AI Framework;
- Ensuring provincial agencies, tribunals and courts are included in the *Framework*;
- Committing to assist municipalities and local agencies develop resources, tools and standards to ensure *Trustworthy AI* in these organizations;
- Developing performance metrics to ensure the province is meeting the goals of Trustworthy AI;
- Establishing a multidisciplinary Trustworthy AI Advisory Group and public consultation plan;
- Committing to meaningful and multidisciplinary public input and participation in all phases of provincial AI regulation and development.

Impact Of AI on Government Administration and Decision-Making

AI and ADM tools will transform government decision-making and administration. This trend will expand and accelerate as AI, machine learning and natural language processing continue to develop.

AI and ADM systems offer significant potential benefits to governments and the public. Many believe that these tools can "crack the code of mass adjudication," improve the accuracy and consistency of government decision-making, improve public services, and reduce government backlogs. Many also believe that government AI systems have the potential to reduce discrimination and enhance democratic and legal accountability.

Governments must respond to the well-documented risks of AI and ADM systems. Experience with government AI and ADM systems across North America, Europe, Australia and New Zealand demonstrates the serious risk of racial bias, "data discrimination" and "black box" decision-making. AI systems also risk reducing judicial and administrative discretion, entrenching automation bias, and undermining the legitimacy of government and agency decision-making. The burden of these harms is not shared equally, as they fall disproportionately on racialized or otherwise vulnerable communities.

Canadian governments have an opportunity to become leaders in successful AI deployment by applying hard-learned lessons from other jurisdictions about how to ensure trustworthy and accountable AI.

Participation

There must be broad participation in the design, development and deployment of government AI systems. Unequal access to information and participation in AI decision-making can worsen existing biases and inequality, result in ineffective public services, and damage trust in government. This participation must include technologists, policymakers, legal professionals and the communities who are likely to be most affected by this technology.

Law Reform

"Techno-utopianism" is risky. There are many examples of government AI systems being implemented too soon and at great risk to many individuals and communities. Systems must be developed thoughtfully, deliberately and incrementally to ensure respect for rights, fairness and transparency.

Comprehensive law reform is needed. The systemic legal issues raised by AI cannot be addressed through individual litigation, best practices or piecemeal legislation. There are many potential legislative or regulatory responses. Choices in this area are complex and consequential.

Accountable AI

The speed, scale and opacity of AI systems will transform the legal and policy landscape governing government decision-making.

Absent proactive measures, AI tools may worsen bias, unfairness and legal accountability in government decision-making. That said, AI tools offer significant potential to improve fairness and enhance legal accountability. Neither outcome is predetermined or inevitable. Whether government AI is harmful or beneficial will depend on choices and decisions made by governments, courts and others in the coming months and years.

Legal accountability for AI-based government decision-making is a fundamentally important dimension of "Trustworthy AI." Legal accountability will depend on a sophisticated mix of technical, operational and legal skills. As a result, a multidisciplinary, multifaceted strategy is needed. AI regulation, by itself, will not ensure legal accountability for government AI decision-making. Legal accountability depends equally on reforms to, or reinterpretations of, existing human rights and administrative law rules; meaningful due process protections; innovative access to justice strategies; and ensuring meaningful opportunities to challenge individual decisions.

Legal accountability strategies are interdependent. Systemic regulation will not provide meaningful accountability unless litigants have a meaningful opportunity to challenge AI-based government decisions and have access to appropriate legal remedies. Similarly, AI litigation is unlikely to ensure legal accountability in the absence of systemic regulation. Individual rights and systemic governance are not in opposition; they complement and support each other.

Government AI legal accountability challenges often combine both legal and technical analysis, a combination that itself raises new questions and difficulties.

Access to Justice

It is possible that only the best resourced and most sophisticated litigants will be able to challenge many AI-based government decisions. Absent proactive initiatives, government AI decision-making may add significant new access to justice barriers to low-income, marginalized, Indigenous and racialized communities, thus compounding the over-representation of these communities in Ontario's justice system.

Human Rights

Government AI systems must be human rights compliant. However, achieving human rights compliance will be difficult unless several important issues are addressed.

Most notably, human rights compliance will depend on how policymakers, courts and tribunals address significant evidential challenges inherent in "black box" government AI systems. These challenges can be addressed, in part, by systemic and significant disclosure, transparency and testing of government AI systems.

Human rights compliance will also depend on thoughtful answers to several equally important legal, technical and practical issues, including:

- Data standards;
- Guidelines or metrics to measure bias and discrimination in AI systems;
- Bias testing or auditing requirements;
- Determining reasonable

accommodations in AI systems;
- Determining remedial provisions sufficient to address AI systematic harms.

Finally, human rights compliance of government AI systems will depend on addressing two further issues:

- Can AI systems be used to reveal or address systemic discrimination?
- Are there AI "no-go" zones where a government AI system's potential risk to human rights is so significant that governments should prohibit the use AI in that area?

Absent appropriate guidance, provincial ministries, agencies, tribunals and/or courts will likely need to address complex legal and technical issues on a case-by-case basis, which may result in poorer public services, inconsistent decision-making, diminished rights protection, delays, added costs and unnecessary litigation.

The provincial government, Ontario Human Rights Commission and others should work together to create new policy guidance to respond to these issues. An important early initiative could be to develop a made-in-Ontario AI Human Rights Impact Assessment to assist developers, policymakers, decision-makers and the public assess the human rights compliance of an AI system.

Administrative Law

Administrative law is likely to have a profound impact on government use of AI, ADM and related technologies. These systems will have to be designed, administered and evaluated to ensure compliance with the principles of procedural fairness and substantive fairness.

The Government of Canada's Automated Decision-making Directive addresses many administrative law issues positively. For example, many features of the Directive raise the standard of administrative governance. The Directive has gaps and shortcomings, however, even within the realm of federal administrative law.

There is no equivalent of the Canada ADM Directive at the provincial level, nor are there equivalent laws, policies or directives at the municipal or provincial agency level. As a result, the challenge of reconciling AI systems and administrative law will be greater in Ontario. Accordingly, the Government of Ontario should adopt a provincial equivalent of the Automated Decision-making Directive.

A provincial Automated Decision-making Directive could address many of the important legal, technical, and practical issues necessary to ensure government AI systems are compliant with administrative law, including but not limited to:

- "Notice" when an AI system is used by the provincial government, municipality or agency;
- How to assess the "reasonableness" of a government decision made or influenced by AI;
- "Reasons" from an AI system;
- "Explainable AI" and legal justifications;
- How to assess risk and impact of a government AI decision;
- Meaningful participation and appeal rights;
- Efficient and cost-effective dispute resolution.

As with human rights issues, the absence of appropriate guidance on AI administrative law issues will likely mean that provincial ministries, agencies, tribunals and/or courts will need to address complex legal and technical issues on a case-by-case basis.

Privacy

Provincial amendments to *FIPPA* and other initiatives in Ontario, Quebec and at the federal level have focused public attention on AI, privacy law and data protection.

Data collection, processing and sharing can raise privacy concerns. A crucial question is how to balance protection of privacy rights, data collection and use, and AI systems that potentially improve access to and quality of healthcare, safety, education, and social services.

As the province moves forward, there are gaps that should be monitored and addressed:

- The effectiveness of de-identification may be limited;
- AI challenges an individual's right to access their personal data;
- Remedies;
- How to ensure privacy and data governance policymaking be transparent and participatory.

Ontarians can learn from the examples in Australia (Robodebt), Michigan (MiDas), and the Netherlands (SyRi). All three systems were legally challenged, widely criticized, and eventually reformed or cancelled due to privacy concerns.

Civil Procedure and Other Issues

Rules of Civil Procedure

Ontario's *Rules of Civil Procedure* will likely allow parties to navigate disputes about artificial intelligence fairly. As the law develops, new AI-specific *Rules of Civil Procedure* should be considered.

Confidentiality and Sealing Orders

The current laws governing confidentiality and sealing orders are likely to be sufficient to address AI-related concerns. Blanket sealing orders, or redaction of confidential information, are likely to be problematic. Other alternatives may be preferable.

Crown Privilege

It is important that the provincial government not be immune to tortious liability for government AI systems. The new provincial *Crown Liability and Proceedings Act*, 2019 raises questions about whether parties can be barred from negligence claims against the provincial government for developing, implementing, deploying and relying on AI systems.

Evidence

"Black box" AI systems could create burdensome and potentially impossible evidentiary thresholds for plaintiffs challenging government AI systems. Evidentiary thresholds in AI cases will be a key issue. The laws of evidence are flexible and adaptable but should be monitored.

Education

It will be necessary to develop training and guidance for participants in Ontario's civil justice system. The provincial government, judiciary, court administrators and provincial legal organizations should develop educational programs and materials for the judiciary, tribunal members, counsel and administrators.

Ongoing Monitoring

The development and use of AI in Ontario's justice system should be monitored. The provincial government, judiciary, academics, NGOs, and legal organizations should consider establishing a working group to analyze, monitor and report on the use of AI and algorithms in Ontario's civil justice system.

A complete list of recommendations is included in Appendix A.

AI In Government Decision-Making

What is AI and ADM?

There is no consensus on the definition of artificial intelligence. AI technology and terminologies are evolving. At present, the definition of AI could potentially include algorithms, ADM, natural language processing, machine learning, neural networks and other technologies.[8]

The Government of Canada's Directive on Automated Decision-making defines AI as "Information technology that performs tasks that would ordinarily require biological brainpower to accomplish, such as making sense of spoken language, learning behaviours or solving problems."[9] This definition is not used consistently, even within Canada.

For the purpose of this paper, the LCO adopts the definition of AI proposed by American commentators David Freeman Engstrom and Daniel Ho:

> *[We] use "artificial intelligence" to mean any instance where an agency deploys models to learn from data with the goal of prediction. AI is thus used interchangeably with machine learning but excludes simple process automation (e.g., a case management system to digitally process benefits applications) and conventional statistical analysis (e.g., regression with the aim of drawing a causal inference).*[10]

Why Are Governments using AI?

The benefits of AI to government decision-making may include increased accuracy, fairness, transparency and efficiency in decision-making.[11] According to a recent NYU/Stanford study,

> *Rapid developments in AI have the potential to reduce the cost of core governance functions, improve the quality of decisions, and unleash the power of administrative data, thereby making government performance more efficient and effective.*[12]

Many commentators believe that "AI might finally help crack the code of mass adjudication, improving accuracy, reducing inconsistency, and cutting down on rampant backlogs that plague [government] agencies."[13] Accordingly, many believe that AI has the potential to significantly improve the administrative state.[14]

There is also a belief among many policymakers, technologists and academics that these tools can make government decision-making fairer, less biased and more equitable than conventional analytic tools and subjective human decision-making.[15] Key to this view is the idea that AI and ADM tools are "harnessing the power of data to aid decision-making."[16]

Finally, many believe AI tools can be used to protect and promote human rights.[17]

Not everyone shares these views. Indeed, AI, algorithms and ADM are often referred to as "weapons of math destruction"[18] or as "a sophisticated form of racial profiling."[19]

Given these experiences, there are many who believe AI systems should never to be used to make or aid government decision-making; that there are no examples of AI being employed to benefit racialized or vulnerable communities; and that AI simply undermines social safety nets in ways that would not otherwise be possible.[20]

Many of the risks and potential harms of AI are discussed in this report.

How Are Governments Using AI?

Governments around the world are using AI across a broad range of areas, services and functions. AI use in government is expanding rapidly, and it is hard to keep track of current applications.

The LCO has surveyed government use of AI systems in the United States to provide readers with an admittedly incomplete sample of where and how the technology is being used. In the US, AI and ADM tools are currently being used to assist government operations "across the full range of governance tasks," including:

- *Enforcing* regulatory mandates centered on market efficiency, workplace safety, health care, and environmental protection;
- *Adjudicating* government benefits, from disability benefits to intellectual property rights;
- *Monitoring and analyzing* risks to public health and safety;
- *Extracting* useable information from the government's massive data streams, from consumer complaints to weather patterns;
- *Communicating* with the public about its rights and obligations as welfare beneficiaries, taxpayers, asylum seekers, and business owners.[21]

The list below is a representative sample of the range and types of AI and algorithmic systems currently being used in the United States. Readers should note that many of these applications are in areas of most concern to access to justice advocates, including adjudication, child welfare, policing and criminal justice, housing, immigration and public benefits.

Adjudication[22]

- Clustering for Micro-specialization. The Social Security Administration (SSA) uses algorithms to cluster similar cases rather than randomly assigning them to adjudicators.
- The Insight Program. The SSA uses natural language processing (NLP) for quality assurance, by improving the quality of decision writing.

Child Welfare[23]

- Child Risk and Safety Assessments are used by child welfare agencies to evaluate potential child neglect and abuse cases for risk of child death/injury.
- Genogram and Ecomap Software is "an assessment tool that allows child welfare caseworkers to map family trees, identify gaps in family history, organize information amassed from family, and assess interventions."

Criminal Justice[24]

- DNA Analysis (probabilistic genotyping) systems are used to interpret forensic DNA sampling by performing statistical analysis on a mixture of DNA from different people to determine the probability that a sample is from a potential suspect.
- Inmate Housing Classifications is "a system that analyzes a variety of criminal justice data and outcomes to determine the conditions of confinement, eligibility for programming, and overall housing arrangements of inmates in a jail or prison."
- Bail, Sentencing and Parole Risk Assessment Tools are "algorithmic systems that use existing criminal justice data to produce a "risk score" to inform decisions made pre-trial, during incarceration, sentencing, and parole/probation."
- Job Training Management Algorithms are "used by the Department of Corrections to determine the amount of compensation that a job training provider should receive as well as who is eligible for enrollment in these programs."

Comment and Complaint Analysis

- Natural Language Processing. The Federal Communications Commission uses natural language processing (NLP) to analyze its comment and complaint webpages regarding the proposed repeal of net neutrality, including sentiment analysis, a component of NLP.[25]
- Natural Language Processing of comments to Consumer Financial Protection Bureau "to automatically analyze text to categorize narratives, identify trends, and predict consumer harm."[26]

Customs and Border Protection

- Automated Targeting Systems use statistical methods to assess particularized future risks.
- Risk Prediction Program "aims to improve the Automated Targeting System in collaboration with the Science & Technology Directorate of the Department of Homeland Security."[27]
- Biometric Recognition including facial recognition technology (FRT).[28]

Education[29]

- Teacher Evaluation Algorithms are used to measure a teacher's impact on student achievement based on standardized test scores.
- School Assignment Algorithms are "used to assign students to schools for K-12 based on preferences, test scores, portfolios, and other criteria," including school district's diversity goals.
- School Violence Risk Assessment tools are used "to identify students who are at a high risk for school related violence (e.g. homicide, suicide)."
- Student Risk Prevention Algorithms predict "students at risk of being arrested or in crisis and agencies", using information including "zip codes, truancy numbers, race, and other indicators. The data used in these systems is often shared with law enforcement and other government agencies."

Fire[30]

- Fire Department Funding Algorithms analyze existing data on fire occurrences, response times, and other variables to inform decisions on funding allocation and closures of neighborhood fire departments within a given jurisdiction.

- Fire Risk Assessments "use data mining to predict which buildings are at highest risk of catching fire."

Food and Drug Administration

- FAERS beta. The Food and Drug Administration's Adverse Event Reporting System Database assists in FDA post-market surveillance.[31]
- NEST. The National Evaluation System for Health Technology system is designed "to 'help improve the quality of real-world evidence that FDA can use to detect emerging safety signals quickly and take appropriate actions."[32]

Healthcare[33]

- Healthcare Delivery and Workflow Decision Systems are software and IT infrastructure intended to provide predictive analytics for care providers and hospital systems to ascertain how best to distribute healthcare resources.
- Healthcare Diversity and Workflow Decision Systems are "intended to provide predictive analytics for care providers and hospital systems to ascertain how best to distribute healthcare resources."

Housing[34]

- Building Inspection Predictive Analytics use public data to identify buildings at the greatest risk for physical deteriorating conditions that endanger the health and safety of residents.
- Tenant Harassment Predictive Analytics "analyze public data to identify landlords with a high likelihood to harass tenants to help prioritize inspections for tenant harassment."
- Homeless Prioritization Algorithms are automated systems (known as coordinated entry systems) that "use information from different government agencies and sometimes third-parties to assess and prioritize allocation of existing housing based on need…"

Immigration[35]

- Investigative Decision System is a software suite that allows for government officials to access an individual's personal and private information such as biometrics, criminal records, work/home addresses and personal connections to identify new targets for deportation and aid in removal proceedings.
- Immigration Detention Risk Assessments are "a computerized system that evaluates an individual's criminal history, work status, likelihood of fleeing and other information to produce recommendation[s] about whether the person should be detained or released prior to a removal hearing."

Policing[36]

- Predictive Policing is "algorithmic software that attempts to identify where a crime may occur in a given window of time in small geographic areas or identify individuals who may be a perpetrator or victim of a crime."
- Automated License Plate Readers are "high speed, computer-controlled camera systems that automatically capture all license plate numbers that come into view, along with the location, data, and time, and sometimes photographs of the vehicle

and its drivers and passengers. Algorithms can be combined with this data to predict where a driver may be in the future."

- Facial Recognition systems are "a computer vision system used to detect faces of individuals and for surveillance purposes, which includes building a database of faces…"
- Social Media Monitoring systems are "a tool […that] analyzes social media messages to identify individuals as part of a gang/crew, find evidence to support an arrest, identify potential crimes."
- Police Mental Health Screening Tools are intended to assist officers "de-escalate interactions with people with mental illness."

Public Benefits[37]

- Public Benefits Fraud Detection Systems are pattern recognition systems used to detect fraud or abuse of public benefits.
- Fugitive Felony Compliance Systems use "an automated program to discontinue an individual's public assistance benefits if they showed up on both the lists of individuals receiving assistance from the federal Supplemental Nutrition Assistance Program (SNAP) and the list of outstanding felony warrants maintained by the law enforcement agency network."
- Medicaid Benefits Algorithms are "used to determine an individual's eligibility for Medicaid, the amount of benefits, compliance, and/or termination."

Public Health[38]

- Disease Surveillance and Treatment Systems are used to identify individuals with or at risk of contracting chronic or infectious diseases for treatment…"
- Prescription Drug Monitoring Programs "mine state prescription drug databases for irregularities that indicate doctor shopping, doctor overprescribing, and other practices that lead to abuse and overdoses."

Securities Regulation

- CIRA. The Securities Exchange Commission's Corporate Issuer Risk Assessment tool was developed to "detect fraud in accounting and financial reporting."[39]
- ARTEMIS and ATLAS. The Advanced Relational Trading Enforcement Metrics Investigation System and Abnormal Trading and Link Analysis System are algorithmic "tools target[ing] trade-based market misconduct."[40]
- Form ADV Fraud Predictor. This system "helps SEC staff predict which financial services professionals may be violating federal securities laws."[41]

Tax Collection and Compliance[42]

- Return Review Program. RRP is a tool being developed by the Internal Revenue Services to "generate risk scores for all national individual tax returns claiming a refund" with the goal of responding to "budget and workforce cuts"

There are many resources identifying additional AI systems being used by governments in the UK, Australia and New Zealand to support decision-making in a broad range of civil, administrative, and criminal justice applications.[43]

Importantly, these applications are the tip of the iceberg of potential government AI applications. Leading US commentator Cary Coglianese states

> *[This] currently limited nature of machine learning will not last long. With the advancement of machine-learning techniques and the proliferation of supporting back-end data infrastructures, the role for algorithms in government is likely to expand. Not only could machine learning soon be employed in more determinative ways, but it could do so, broadly speaking, to yield two different kinds of determinations: adjudications and regulations.*
>
> *For adjudication by algorithm, no longer might algorithms merely inform adjudicatory decisions, such as by targeting inspectors to certain facilities or flagging tax returns for a full review by human auditors. Rather, machine learning, in conjunction with other computer systems, might directly and automatically conduct an audit or inspection, deem a tax return fraudulent, decide whether an individual should receive an airplane pilot's license, award or withhold disability benefits, or assign prisoners to cells based on predictions of their propensity for future violence. It takes little technical imagination to see how these applications could materialize; they would be relatively straightforward applications of machine learning.*[44]

The potential reach of AI and algorithms to aid justice system decision-making is staggering. The AI Now Institute, a leading American AI research organization, notes

> *...[a]utomated decision systems can exist in any context where government bodies or agencies evaluate people or cases, allocate scarce resources, focus scrutiny or surveillance on communities, or make nearly any sort of decision.* [45]

Transposed to the Canadian context, the applications currently in use internationally would affect some of Canadian's most important government services and the jurisdiction and workload of many Superior Courts, provincial courts, administrative tribunals, government ministries, agencies and municipalities.

Growing use of AI and algorithms has been found in government-use studies in the US, UK, Australia and New Zealand.[46] Unfortunately, at present there are no equivalent Canadian studies. Nor is there a central list or repository of AI or automated decision-making systems in use in the Canadian justice system or other Canadian government applications.[47] As a result, it is very difficult to assess how widespread this technology is being used in Canada. Disclosure of AI systems used by Canadian governments and public agencies is discussed extensively in the LCO's *Regulating AI* paper.[48]

How AI Transforms Government Decision-making

The LCO believes that the shift from human decision-making to AI or "hybrid human/AI" systems fundamentally transforms the legal and policy landscape of government decision-making. For example, AI embeds a complex mix of legal, technical, statistical and operational decisions into code. AI systems also rely on data to support every stage of AI decision-making. These changes, and others, raise significant new questions about how to ensure legal accountability for government decision-making.

Most discussions about legal accountability and government AI systems focus on issues of bias and "black box" decision-making. This focus is understandable but incomplete, as the transition to government AI decision-making raises many additional challenges to legal accountability.

What follows below is a high-level summary of how AI can change government decision-making and the resulting challenges for legal accountability. Readers should note that many, if not most, of these changes are contested and controversial. AI proponents identify the benefits of many of these changes; critics identify the risks.

- **AI accelerates the speed and scale of government decision-making**.[49] AI systems can make decision/predictions much faster than humans. As a result, some commentators believe that "AI might finally help crack the code of mass adjudication… cutting down on rampant backlogs that plague [government] agencies."[50] On the other hand, the speed and scale of AI decision-making can mean that errors are magnified and have abrupt, widespread and significant implications.[51] The legal system will need to respond to the speed and scale of AI decision-making.
- **AI may (or may not) improve the objectivity, accuracy and consistency of government decision-making.** Many commentators believe that AI systems are – or can be – neutral, consistent and evidence-based tools that transform arbitrary, opaque, inconsistent and often-biased human decision-making. In this view, AI tools are seen as potentially superior to human decision-making in that AI may "eliminate the variability, indeterminacy, and apparent randomness—indeed, the subjectivity—of human prediction."[52] Critics doubt that AI decisions are necessarily more accurate or consistent, particularly if those decisions reflect bias or if an AI system does not accurately reflect law or policy.[53] The legal system will need to be able to test the objectivity, accuracy and consistency of AI decision-making.
- **AI systems can be biased.** One of the original and compelling justifications for introducing AI and algorithmic systems in government is their purported ability to reduce or at least identify bias.[54] Critics disagree, noting that many AI and algorithms are created using historically racist, discriminatory or biased data.[55] Notably, biased datasets are just one form of potential AI bias. Accordingly, the legal system will need to develop new mechanisms and tests to evaluate government AI systems for bias.
- **"Black box" decision-making/Embedding decisions into code.** AI systems embed an extraordinarily complex mix of legal, technical, statistical and operational decisions into code. The complexity and opacity of AI tools may make AI-aided government decisions "even more inscrutable than human judgments."[56] As a result, even simple algorithmic and AI systems can be complex and opaque "black boxes." But the opposite might also be true: Many believe that, in the long run, AI may formalize and make explicit government policies and choices, thus rendering government decision-making relatively more transparent.[57] The legal system's ability to disclose, understand and evaluate millions of lines of code will be a necessity to ensure government AI legal accountability.
- **Data issues and choices are pervasive in AI systems.** Data trains an AI system, data calibrates an AI system, and the output of AI system is often some kind of "score" or statistical prediction. Data issues and choices can be obscure, consequential and controversial.[58] For example, many AI systems and algorithms are criticized as being racist or based on biased data. Other data issues include

questions regarding statistical "metrics of fairness" and the accuracy, reliability and validity of datasets. The legal system will need to ensure its ability to review and evaluate data issues to ensure legal accountability of government AI systems.

- **AI systems can obscure or blur the distinction between predictions and policy.** A government AI risk assessment tool (such as a bail or sentencing algorithm) is a statistical tool that makes predictions. Program administrators or AI developers will often decide how to turn those predictions into "action directives" or legal decisions. In operation, the distinction between statistical predictions and policy decisions can be blurred or misunderstood.
- **AI systems may change the decision-maker of government decisions.** Instead of having a human decision-maker, AI decision-making shifts many policy decisions to the designers of an AI system. In some cases, decisions are shifted into the "black box" of the AI system. Moreover, the operation and outputs of more advanced AI systems may not be understandable by the system's designers.[59] In some situations, shifting a government decision from a human to a machine could raise concerns about whether the governmental department is improperly delegating its duty.
- **AI may reduce discretion and decision-making independence, even when there is a "human in the loop."** AI can rapidly accelerate and standardize government decision-making, bringing many benefits. At the same time, however, automation bias and related factors may mean that governments and other decision-makers (such as courts or tribunals) intentionally or inadvertently limit their discretion by relying too heavily on machine outputs.[60] Similarly, over-reliance on AI predictions or recommendations may compromise the appearance or reality of judicial or tribunal independence.
- **AI systems can reduce public engagement in policymaking and oversight of important government services and functions.** AI is a complex and evolving set of technologies. Experience in several jurisdictions, including New York City, demonstrate that public engagement on AI policy issues depends on ensuring members of the public are educated and supported before, during and after the AI policy-making process.[61]

It is important to note that the LCO is not suggesting that human decision-makers are necessarily better than AI-based decision-making. On the contrary, much of the enthusiasm for AI in government is based on an appropriate acknowledgement of the inefficiency, inconsistency, and unfairness of human decision-making. Notwithstanding this acknowledgement, the potential benefits of AI decision-making cannot overshadow need to ensure AI decision-making is legally accountable.

The next section considers the concept of "Trustworthy AI" and how Ontario's policymakers have responded to the challenges of AI.

Ontario's Trustworthy AI Framework

Introduction

Many governments have responded to the challenges of AI by announcing or adopting "trustworthy AI" or "ethical AI" frameworks and policies.[62] Generally speaking, these frameworks are designed to assure the public and stakeholders that government AI development and use will be transparent, legal and beneficial. The critical reception to many of strategies has been mixed. Many strategies have been harshly criticized as "whitewashing biased tech,"[63] "dysfunctional" or "not fully addressing the myriad harms that may arise from the opacity, complexity, scale and power-imbalance in which AI systems are deployed."[64]

The Government of Canada's Directive on Automated Decision-making ("the Canada ADM Directive") is the most notable Canadian "trustworthy AI" framework. Internationally, the most comprehensive and significant "trustworthy AI" framework is the European Commission's proposed AI rules ("EC AI Proposal").[65] The LCO and the Research Chair on Accountable Artificial Intelligence in a Global Context recently published a paper comparing the Canada ADM Directive and the EC AI Proposal.[66]

As part of their *Digital and Data Strategy*, the Government of Ontario has been working to develop its own *Trustworthy AI Framework*.[67] The LCO commends the provincial government for its commitment to "trustworthy AI." In this section, the LCO briefly reviews steps taken by the provincial government to date and concludes there is still much work to be done.

In May 2021, the province commenced a consultation process asking the public to provide input and ideas on how government can develop an AI Framework that is "accountable, safe and rights based."[68] The province framed the consultation through three key commitments: (i) No AI in secret; (ii) AI use Ontarians can trust; (iii) AI that serves all Ontarians. The consultation process included the release of the province's draft Alpha documents on *Principles for Ethical Use of AI* and *Transparency Guidelines for AI*.[69]

The LCO provided commentary on the Alpha Documents in June 2021.[70]

Following email submissions and an online survey, the province produced a "What We Heard" document in July 2021.

In September 2021, the provincial government announced its "Action Plan."[71] The Action Plan had two key components. The first was a commitment to operationalize the *Principles for Ethical Use of AI, and Transparency Guidelines for AI* (the "Alpha documents"). The second was to improve the province's AI and Algorithm Inventory.[72] The Action Plan was broad enough to give province significant leeway, but included an emphasis on public consultations.

The provincial government continued targeted consultations and in January 2022 released its "Beta principles for the ethical use of AI and data enhanced technologies in Ontario" (the "Beta Principles").[73] The Beta Principles include six points to "align the use of data enhanced technologies within government processes, programs and services with ethical

considerations and values." The Beta Principles stated that provincial AI systems should be:

- Transparent and explainable;
- Good and fair;
- Safe;
- Accountable and responsible;
- Human centric;
- Sensible and appropriate.

The Beta Principles appear comprehensive. They highlight the most significant issues in AI design and operation and address many of the concerns mentioned in the Trustworthy AI consultation process. For example, the Beta Principles:

- Acknowledging the need for transparency, the Beta Principles state that an explanation provided about the AI system must be meaningful to the person requesting it.
- State that AI systems must respect the rule of law, human rights, civil liberties and democratic values, which include dignity, autonomy, privacy, data protection, non-discrimination, equality and fairness.
- State that there must be a public and accessible process for redress that is developed with input from a multidisciplinary team and multiple stakeholders.
- Acknowledge that AI systems should be peer reviewed or audited to ensure that unwanted biases have not crept in over time.

Importantly, the Beta Principles highlight the importance of these considerations throughout the life cycle of the AI system.

The Beta Principles are an important step in the process to develop an AI Framework. As of May 2022, however, the provincial government's next steps are unclear.

Analysis

As noted above, the LCO provided comments on the province's Trustworthy AI Alpha documents in June 2021. Most of the recommendations in that submission remain apt, and the LCO reiterates them here with some modifications. These recommendations crystalize key features of the LCO's analysis over the course of our work in this area.

In the LCO's view, these recommendations are necessary first steps that would provide Ontarians with important assurances that the provincial government is dedicated to thoughtful, transparent, accountable and bias-free AI systems.

In order to support Trustworthy AI in Ontario, the LCO makes the following recommendations:

Recommendation 1

The provincial government should not deploy high-risk AI or automated decision-making technologies prior to adoption of its comprehensive Trustworthy AI Framework.

Recommendation 2

The Trustworthy AI Framework should be established in legislation and regulations.

As noted in the LCO's *Regulating AI* report, there is an ongoing international discussion regarding which AI systems should be prohibited or, at a minimum, categorized as high risk and therefore subject to more extensive regulatory requirements.[74]

It would severely compromise the purpose and credibility of any provincial Trustworthy AI Framework if the province were to deploy a high-risk AI system *prior* to the adoption of its comprehensive Trustworthy AI Framework. Similarly, high-risk systems currently in operation or under development should not be "grandfathered" or exempt from the provincial Trustworthy AI Framework. Risk is risk, irrespective of whether an AI system was developed/deployed before or after an arbitrary proclamation date.

The LCO further recommends the provincial government consult broadly with stakeholders to identify prohibited and/or high-risk AI applications. These consultations could be organized or evaluated by the Trustworthy AI Expert Advisory Task Force recommended below. As a preliminary matter, provincial policymakers and the Task Force should seriously consider whether systems that target or disproportionately impact or harm vulnerable populations such as disabled persons, children, Indigenous or racialized communities and/or low-income communities should be preemptively prohibited or identified as high-risk. AI and related systems have significant potential to affect the legal rights (including but not limited to human rights) of these communities. As a result, there are good reasons for these systems to be subject to prohibitions or higher regulatory standards.[75]

Facial recognition and biometric identification technologies are notably high-risk.

Legislation is necessary to provide the foundational governance framework for these systems. A legislative framework would provide consistent direction and accountability requirements to the actors, departments and agencies within its scope. It would also ensure changes to the governance framework were subject to legislative and public review. Finally, legislation would establish a level of public and legal accountability commensurate with the issues and rights at stake.

The proven risks of AI cannot be comprehensively addressed through individual litigation, best practices, existing or piecemeal legislation. Law reform is needed to ensure AI and ADM systems meet high legal standards regarding disclosure, legal accountability, equality, procedural fairness/due process and access to remedies. The appropriate elements of the LCO's recommended legislative provisions are set out in Recommendations 10 and 11 and the LCO's *Regulating AI* paper.[76]

Ethical AI guidelines are insufficient to mitigate the harms caused by the use of AI and related systems due to their lack of specificity and reliance on voluntary compliance. Ethical guidelines, directives, "playbooks" or best practices and other "soft law" instruments have potential to *supplement* mandatory legal obligations and requirements, but they are not a substitute.[77]

The LCO is concerned about early signs that the provincial government may be considering adopting an "ethical AI" voluntary code of compliance or best practices. The provincial government's repeated references to "guidance" and "principles" may signal this approach. If so, the province's Trustworthy AI Framework may not meet the risks of AI systems and the demands of AI governance/regulation.

The issue of "soft" verses "hard" rules and regulations featured prominently in provincial consultations and the LCO's *Regulating AI* report. "Soft law" practices and guidelines are useful tools to address some AI-related issues.[78] They are not sufficient to protect human rights and other important legal rights.

Recommendation 3

The Trustworthy AI Framework should promote AI transparency, accountability, and public engagement in the development, operation, and evaluation of provincial AI systems.

The province's commitments take important first steps to fulfilling Recommendation 3. The LCO believes, however, that the province needs to fulfill all the recommendations in this report to meet this goal.

Recommendation 4

The provincial government should create an AI framework to specifically address AI systems that are developed, or used in, the criminal justice system, such as facial recognition, biometric identification, predictive policing and bail/sentencing risk assessments.

The LCO's *The Rise and Fall of Algorithms in the American Justice System: Lessons for Canada* report discusses the risks of AI and automated decision-making systems in the criminal justice system at length.[79] These risks include, but are not limited to, *Charter* violations, biased data, the "metrics of fairness", data transparency and opacity, "data scoring", algorithmic bias, lack of due process and a lack of access to justice.

Unfortunately, the Canada ADM Directive does not include AI or automated decision-making systems in the federal criminal justice system. The European Commission proposed AI rules include detailed provisions identifying AI systems in "law enforcement" as being preemptively high-risk, and thus subject to more detailed and expansive regulatory requirements.[80]

Ontario should commit to developing a dedicated Trustworthy AI Framework to address AI tools used in the criminal justice system. These tools could include facial recognition and other biometric assessments within its scope. The deployment of any AI tools in the criminal justice system in Ontario should be delayed until a Trustworthy AI framework has been developed. AI in the criminal justice system requires unique attention because of the fundamental rights, liberty interests, and unique statutory and procedural issues affected. Equally important, a dedicated criminal justice AI Framework will need to acknowledge the over-representation of low income, Indigenous and marginalized communities communities and address the.potential for human rights abuses/adverse impact on historically disadvantaged communities in Ontario.

Recommendation 5

The Trustworthy AI Framework should establish a framework for municipalities, provincial agencies, and courts and tribunals under provincial jurisdiction.

Recommendation 6

The provincial government should commit to assisting municipalities and public agencies develop resources, tools, and standards to ensure Trustworthy AI in these organizations.

Recommendations 5 and 6 are related. At present, there is a significant regulatory "gap" in Canada. The Canada ADM Directive guides AI and automated decision-making development for most federal departments and agencies. Unfortunately, there are no equivalent regulatory instruments governing public sector AI systems under provincial jurisdiction. This is why the provincial government's Trustworthy AI Framework is so important.

There is a significant risk that the breadth and impact of Ontario's Trustworthy AI Framework will be undermined or blunted if its reach does not extend to all governments, agencies, courts and tribunals within provincial jurisdiction. Experience in other jurisdictions proves that some of the most consequential and risky AI systems have been deployed by municipalities or local institutions, such as child welfare agencies and police services.[81] The province should also work collaboratively with Indigenous governments and agencies to accomplish this goal consistent with reconciliation commitments.

Ontario's Trustworthy AI Framework should close these gaps. Consistency in standards across provincial jurisdiction would allow all Ontarians to benefit from the protection of the Framework and encourage trustworthy AI innovation and development across the province.

Recommendation 7

The provincial government should develop public performance metrics to ensure the province is meeting the goals of Trustworthy AI.

It will be important to track the progress and "success" of the provincial Trustworthy AI Framework through performance metrics. These metrics should focus the provincial government's reform efforts and provide necessary public transparency and accountability.

Recommendation 8

The provincial government should establish a multidisciplinary Trustworthy AI Expert Advisory Task Force and public consultation plan to advise provincial policymakers on how to fulfill the commitments and recommendations herein.

AI regulation is a complex undertaking, involving multidisciplinary stakeholders and the thoughtful balancing of complicated rights and objectives. Given the pace of AI adoption across Ontario, Canada and internationally, there is a need for the provincial government to act to deliberately and efficiently.

In order to meet these priorities, the provincial government should establish a multidisciplinary Trustworthy AI Expert Advisory Task Force to advise provincial policymakers on how to fulfill the commitments identified in this submission and enshrine them in provincial legislation/regulations.

The LCO believes the starting point for AI regulation is robust and ongoing public participation. More specifically, the LCO believes that governments must engage with technologists, policymakers, government managers, frontline staff, lawyers, industry associations, community organizations and, crucially, the stakeholders and communities who are likely to be most affected by this technology.

The LCO emphasizes that communities (including Indigenous, racialized or otherwise marginalized communities) may be better positioned than lawyers, academics, advocates or regulators to identify some of the risks *and* benefits of AI and related technologies. These communities have both experience and expertise that is crucial to thoughtful regulation.

Recommendation 9

The provincial government should continue to seek meaningful and multidisciplinary public input and participation in all phases of AI regulation development.

The LCO believes it is essential the public be invited to provide input before, during and after the initial development and regulation of AI and related systems. Proactive participation is likely to promote good governance, thoughtful regulations and engender public trust in public sector use of AI and related systems.

Untrustworthy AI

There are many examples of "untrustworthy" government AI and ADM systems in the civil and administrative justice systems. A small sample of noteworthy and high-profile government AI systems that have been harshly criticized, paused and/or withdrawn includes:

- **A-Level Grading Algorithm (United Kingdom)**, an algorithm used to estimate students' A-level and GCSE grades based on the historic performance of individual secondary schools.[82]
- **Allegheny Family Screening Tool (United States)**, a predictive risk algorithm that analyzes allegations of child mistreatment.[83]
- **Arkansas Medicaid Benefits (United States)**, an AI system to determine how many hours of homecare a person was entitled to from Medicaid.[84]
- **Clearview AI (Canada)**, a facial recognition system and database used by the RCMP and other police services.[85]
- **COMPAS (United States)**, a risk assessment algorithm used to predict recidivism in bail and sentencing proceedings.[86]
- **Houston Educational Value-Added Assessment System (EVAAS) (United States)**, an algorithm that measured teacher effectiveness by tracking the teacher's impact on student test scores over time.[87]
- **Michigan Integrated Data Automated System (MIDAS) (United States)**, an unemployment insurance algorithm designed to identify potential fraud.[88]
- **PredPol (United States)**, an algorithm designed to predict the location or perpetrators of crime.[89]
- **Online Compliance Intervention System ("Robodebt") (Australia)**, an automated debt recovery system designed to recover income support overpayments.[90]
- **SyRi (The Netherlands)**, an algorithm designed to identify individuals who might be committing social benefits or tax fraud.[91]

These examples, and others, illustrate the risks and harms of poorly designed, ineffective, unfair, biased, or otherwise unaccountable government AI systems.

Government AI system developers, administrators, policymakers, and stakeholders can learn important lessons from these examples about "untrustworthy" AI systems. These examples also illustrate how and why AI and ADM systems tend to be challenged, including how legal challenges are often combined with media or political pressure. The balance of this section considers emerging research on "cancelled" AI and ADM systems and analyzes a small sample of case studies to illustrate many of the challenges and issues discussed in this report.

Cancelled Government AI and ADM Systems

A recent report by the Data Justice Lab, a research institute at the University of Cardiff, investigates why government departments and agencies in different countries decide to pause or cancel their use of automated decision-making systems.[92] This report, *Automating Public Services: Learning from Cancelled Systems,* is one of the first systemic reviews of "cancelled" government AI and ADM systems.[93]

The Data Justice Lab reviewed 61 government ADM systems that have been cancelled or paused. These systems were in Europe, the United Kingdom, North America, and Australia/New Zealand. Some of the key findings in the report include:

- Systems were most often paused or cancelled in the justice sector (policing and law) (32), followed by systems used to detect fraud in government benefits programs (12), systems used to predict risk in child protection (5), systems used in education (4), and immigration (3). The balance of cancelled systems were in the finance, border control, urban planning, and health sector.
- Systems were paused or cancelled for a variety of related and often overlapping reasons, including concerns about a system's accuracy and effectiveness; civil society critique or protest; critical media stories; legal proceedings; criticisms that a system violated privacy or due process rights; criticisms that systems were biased or racist; critical government or oversight agency reviews; and governance issues relating to procurement and ownership.
- Although systems were paused or cancelled at different stages of development (for example with some at the investigation stage, or after a pilot was completed) over half were cancelled after the system was implemented or in use.

The Data Justice Lab report highlights how negative media coverage, independent reviews, and legal challenges often combine to lead governments to pause or cancel systems. The authors state that

> *…researching the factors and rationales leading to cancellation provides a means to get beyond the myths of the technology to better understand its limits and acceptability…This understanding also highlights some of the precautionary steps that may be necessary for governments to make informed decisions.*[94]

The following examples illustrate some of these general findings.

Arkansas Medicaid Benefits

In 2016, the Arkansas government started using an AI system to determine how many hours of homecare a person was entitled to from Medicaid.[95] Under the previous regime, a registered nurse (RN) would come to an applicant's house and personally assess a client's need for homecare. The RN could adjust the hours of eligible homecare based on his or her in-person assessment.

Under a new algorithmic system, the RN would attend an applicant's house and ask a series of fixed questions. The answers to the questions were input into the algorithm, which would then determine the number of hours the client was eligible for.[96] The RN had no discretion to adjust hours allotted by the algorithm.[97]After the algorithm was implemented, approximately 4,000 recipients noticed their homecare benefits had been cut, even though there was no change in their circumstances.[98] Applicants were not given any explanation. In many cases, the RNs simply told them "The computer did it".[99]

Legal Aid of Arkansas filed state and federal lawsuits challenging the algorithm based on constitutional due process and fair notice. Their basic argument was that the algorithm was so complicated that it could not be contested.[100]

The State of Arkansas released the algorithm's source code, which was approximately twenty pages long.[101] Legal Aid of Arkansas analyzed the source code and determined that software errors inappropriately reduced benefits for applicants with cerebral palsy and diabetes.[102]

One of Legal Aid of Arkansas' key concerns was that the government could not understand or figure out if the software was working properly. Legal Aid of Arkansas further argued that they did not have sufficient information about the "black box".[103] In response, the State of Arkansas argued the system was intended to rationalize decisions and make assessments consistent and objective, not to cut benefits.[104]

In the end, a US Federal Court decided that the State of Arkansas did not have to explain how the *system* worked, but that the state agency had to provide *individuals* with an explanation.[105] This relief was of limited benefit, as it could not provide relief to every benefit recipient affected by the new algorithm.[106] Nevertheless, the publicity surrounding the controversy led the State of Arkansas to abandon the original algorithm and replaced with a new and somewhat improved system.[107]

Houston Teacher Evaluations

In 2011, the Houston School District introduced an initiative to measure teacher effectiveness by tracking the teacher's impact on student test scores over time. This initiative, called the Educational Value-Added Assessment System (EVAAS), was developed by SAS, a private technology firm.[108] The EVAAS score became a significant part of teacher evaluation in Houston.

The Houston Teachers Union brought a constitutional claim in US Federal Court challenging the EVASS system on three grounds: procedural due process, substantive due process, and equal protection.[109] This case illustrates some of the procedural and constitutional issues facing litigants wanting to challenge government AI systems.

Procedural Due Process

The Teachers Union argued that there was a "lack of sufficient information to meaningfully challenge terminations based on low EVAAS scores." Counsel for the teachers referred to this as the "black box argument."[110] They argued that the system was "highly secretive and impossible to replicate."[111]

In response, SAS refused to disclose the system's source code, format, training, or data due to concerns about "trade secrets" and fear that other companies would compete with them.[112] The defendant school board argued that the "the Due Process Clause does not empower Plaintiffs to put SAS out of business" by requiring disclosure of its trade secrets.[113]

The Court held that if the school board was going to use the SAS system to terminate teachers, it had to provide teachers with some level of procedural due process, including giving teachers an explanation or access to why and what happened.[114]

Substantive Due Process

The Teachers Union also raised two substantive due process claims:

First, the union argued that there was "no rational relationship" between EVAAS scores and the school board goal of employing effective teachers.[115]

Second, the union argued that the evaluations were too vague, meaning that teachers did not understand what was going on or what they needed to do to improve.[116]

The union was unsuccessful on both arguments. The Court held that to succeed on the first ground, the Teachers Union had to prove that the EVAAS system is not "a rational means of advancing" the school board's goal of "having an effective teacher in every classroom so that every student is set up for success."[117] The Court granted the Board summary judgment on this issue saying "Even accepting the [teachers'] criticisms at face value, the loose constitutional standard of rationality allows governments to use blunt tools which may produce only marginal results."[118]

The Teachers Union also failed on the second substantive due process argument. The Court held that the Teachers Union did not prove that the EVAAS system was "unconstitutionally vague."[119] The Court found that teachers were given general information about the EVAAS system, and a system can be "unfair" or "prone to error", but still not be vague.

Equal Protection

Finally, the Teachers Union argued that the system violated the US Constitution's Equal Protection Clause.[120] In order to do so, the union had to prove that

> *(a) state actor intentionally discriminated against [him] because of membership in a protected class, or (b) he has been intentionally treated differently from others similarly situated and that there is no rational basis for the difference in treatment.*[121]

The Teachers Union argued that the school board wrongly classified teachers based on their EVAAS scores and subjecting them to different treatment based on their scores.[122] The Court disagreed, finding that the EVAAS scores was not a classification system.

In the end, the Teacher's Union defeated the Houston School District's summary judgment motion on procedural due process grounds. Following the unsuccessful summary judgment motion, the School District abandoned the algorithmic system due to political and community pressure: Teachers were leaving the Houston School District because they did not believe the EVAAS was reflective of their teaching abilities. The departure of effective, well-liked teachers caused enough upset in the community that the school district scrapped the system.

The Netherlands Benefits Fraud

Welfare Fraud

In 2014, the Dutch government introduced an AI system called SyRi (System Risk Indication) to identify individuals who might be committing social benefits fraud and tax fraud.[123] Three SyRi projects targeting specific neighbourhoods in the Netherlands were launched to analyze and cross-reference data from a variety of government departments previously kept separate – such as employment, personal debt, taxes, permits, benefit records, and education and housing histories.[124] Its objective was to find "unlikely citizen profiles" that warranted further investigation. The system created a mathematical risk model to identify individuals who were suspicious. For example, SyRi compared social welfare data with home water usage to search for discrepancies between the number of people reported living in a household and the household's water usage. In SyRi's first phase, all residents of certain neighbourhoods were investigated. Individuals considered "high risk" were flagged for further investigation.[125]

The system was designed with certain checks in place, but these proved to be insufficient. For example, the government provided notice of the program, but the information was vague and provided through a City Bulletin not read by populations targeted by SyRi. Citizens could find out if their names were on the "risk indication" register, but were not told the reasons they had been flagged.

SyRi became a major public controversy in the Netherlands. The breadth and impact of the system raised many questions about government surveillance, privacy, discrimination, and "black box" AI and algorithms. For example, Philip Alston, United Nations Special Rapporteur on extreme poverty and human rights, stated that

> *Whole neighbourhoods are deemed suspect and are made subject to special scrutiny, which is the digital equivalent of fraud inspectors knocking on every door in a certain area and looking at each person's records in an attempt to identify cases of fraud, while no such scrutiny is applied to those living in better off areas.*[126]

SyRi was challenged in Dutch courts on several, related grounds, including lack of notice, privacy, lack of explanation, discrimination, and the right to social security. The Dutch government refused to provide any information about how the system worked or the source code for fear that individuals could "adapt their behaviour."[127] As a result, no one could know what data was used, the indicators to predict risk, or the innerworkings of the algorithm that produces the risk score.[128] Further, it was revealed that SyRi was applied primarily against underprivileged, immigrant and vulnerable populations.[129]

In late 2019, a Dutch court found that the government had not found the proper balance between government interest in detecting fraud and protecting individual privacy rights. The lack of transparency was a significant factor in the court's decision.[130] As a result, SyRi was found to be in contravention of human rights and privacy laws.[131]

Childcare Benefits Fraud

There was further AI controversy in the Netherlands when it was discovered that thousands of parents and caregivers were falsely accused of childcare benefit fraud by the Dutch tax authorities.[132] The scandal led to the fall of the Dutch Cabinet in 2021.[133]

In 2013, an algorithm to detect fraud in childcare benefit applications was introduced. The system did not account for insignificant errors or good faith mistakes made by applicants which led thousands of families mistakenly being identified as fraudulent.[134] Families flagged by the system who were considered "at risk" of committing fraud had their benefits suspended, and had to repay large sums of money immediately.[135]

To make matters worse, a factor considered by the algorithm was whether an applicant was a Dutch citizen, and non-Dutch applicants received higher risk scores. The Dutch Data Protection Authority determined that use of citizenship as a factor in childcare allowance applications was discriminatory, unnecessary and inappropriate.[136] Amnesty International reported that the use of citizenship as a factor amounted to racial profiling:

> *The use of nationality in the risk classification model reveals the assumptions held by the designer, developer and/ or user of the system that people of certain nationalities would be more likely to commit fraud or crime than people of other nationalities. It is also indicative of the tax authorities' perception that there is a link between race/ethnicity and crime, as well as an acceptance of the practice of generalizing the behaviour of some individuals to all others who are perceived to share the same race or ethnicity.*[137]

Discovery of the system led to interrogation of parliamentarians to find out whether ministers were aware of the program.[138] Investigations concluded that the Rule of Law had been violated and the judiciary, the cabinet and the House of Representatives were responsible.[139] The controversy led to the fall of the Dutch Cabinet in 2021.[140]

AI Accountability Through Regulation

The LCO and other organizations have proposed many different legal tools or strategies to ensure legal accountability for government AI systems.[141] These proposals respond, in part, to the hard lessons learned with many early government AI systems, including lessons about bias, "black box" decision-making, lack of participation, inaccurate systems and secrecy. In so doing, these proposals are designed to ensure – or at least facilitate – trustworthy AI systems that promote the benefits of AI while minimize its harms.

In August 2021, the ADA Lovelace Institute, the AI Now Institute and the Open Government Partnership produced a report summarizing the lessons learned from what they describe as the "first wave" of "algorithmic accountability policy for the public sector."[142] This report, titled *Algorithmic Accountability for the Public Sector,* identified eight policy mechanisms "through which governments have sought to achieve algorithmic accountability in the public sector," including principles and guidelines, prohibitions and moratoria, public transparency, impact assessments, audits and regulatory inspection, external/independent oversight bodies, rights to hearings and appeals, and procurement conditions.[143]

The LCO's April 2021 Issue Paper, *Regulating AI: Critical Issues and Choices*, set out the first comprehensive framework for regulating government AI systems in Canada. That paper recommends a comprehensive cascading "mixed model" of regulations, practices and standards the LCO believes are needed to address the well-established risks and potential harms of government AI systems.[144] The LCO report provided a uniquely Canadian analysis of the policy instruments identified in the *Algorithmic Accountability in the Public Sector* report.

As noted in the LCO's *Regulating AI* report, the LCO believes there is an urgency to developing thoughtful AI regulations, as growing public anxiety (and controversies) about AI and ADM in areas as diverse as criminal justice, benefits determination, education, employment, surveillance, consumer profiling, and even urban design have led to public expectations about the need for "trustworthy" AI.

Readers are encouraged to review the LCO's *Regulating AI* report for more detailed analysis of these issues. For present purposes, however, the LCO will reiterate its earlier recommendations as we believe these are a key component of "accountable AI."

In order to support accountable AI in Ontario, the LCO makes the following recommendations:

Recommendation 10

The provincial government's Trustworthy AI Framework should be established in legislation and regulations. The legislation should include, but not be limited to, provisions to ensure provincial AI, ADM and related systems are transparent, accountable, and legal. Legislation should also include provisions that promote access to justice, address bias/discrimination, and a requirement to mitigate harms.

Recommendation 11

The comprehensive regulatory regime should include:

- **Baseline requirements for all public sector AI, ADM and related systems, irrespective of risk.**
- **Strong protections for AI and ADM transparency, including disclosure of both the existence of a system and a broad range of data, tools and processes used by the system.**
- **Mandatory "AI Registers".**
- **Mandatory, detailed and transparent AI or algorithmic impact assessments, including the identification of prohibited and high-risk systems.**
- **Explicit compliance with the Charter, human rights legislation and administrative law.**
- **Explicit requirements to measure, correct and audit/monitor bias in AI systems.**
- **Data standards.**
- **Access to meaningful remedies.**
- **Mandatory auditing and evaluation requirements.**
- **Independent oversight of both individual systems and government use of AI, ADM and related systems generally.**

Both the LCO and the authors of the *Algorithmic Accountability for the Public Sector* report emphasize that AI accountability tools are "nascent, context-specific and that there is no silver-bullet solution to hold algorithmic systems accountable."[145] As in many areas of public policy, it will be important for policymakers to commit to updating and evaluating the effectiveness of new AI accountability measures.

Why regulate AI?

- Government use of AI and ADM is expanding rapidly.
- AI and ADM systems can raise significant, novel and systemic risks to human rights, procedural fairness and access to justice, including:
  - Risk that AI and ADM systems are racist and discriminatory in their design or outcomes.
  - Risk that "black box" systems obscure legal decisions and choices.
  - Risk that systems violate procedural fairness, disclosure, notice, transparency, explainability and remedy requirements.
- Regulation is needed to ensure systems are transparent, explainable, accountable and comply with *Charter*, human rights and administrative law principles.
- Regulation supports AI innovation and development by promoting trustworthiness, fairness and legitimacy of government and justice-system decision-making.
- Regulation promotes better public services and provides guidance and support to many organizations and individuals considering or actively developing AI and ADM systems.

Accountable AI Through Litigation

The previous section considered regulation as a means to promote AI accountability. This section considers "regulation by litigation" issues, including the emerging right to contest AI and related access to justice issues.

The Right to Contest AI

Government AI regulation proposals are designed to provide a systemic framework for government AI decision-making. These proposals typically do not provide rights to challenge an AI tool in individual cases.

The "right to contest" AI is an emerging AI accountability strategy that would enshrine an explicit, individual right to contest AI decisions. American scholars Margot Kaminski and Jennifer Urban discuss this concept in a recent article.[146] Kaminski and Urban note that a freestanding right to contest AI has been "largely ignored" in the United States in favor of system-wide regulation.[147]

In contrast, Canadian analysts and policymakers have considered ways of enshrining "contestability" rights in government AI regulations. For example, in November 2020, the Office of the Privacy Commissioner of Canada released a report titled, *A Regulatory Framework for AI: Recommendations for PIPEDA Reform,* which included recommendations that

> *…individuals should be provided with a right to contest automated decisions. This would apply both to those scenarios where an individual has provided consent for the processing of their personal information as well as those where an exception to consent was used by the organization. It serves as a complement to the right to explanation.*
>
> *This right would take a similar approach to Article 22(3) under the GDPR, where in certain circumstances an individual can express their point of view to a human intervener, and contest the decision, except it would not be limited to decisions based "solely" on automated processing as in the GDPR. This ability would reduce the risk of algorithmic discrimination or other unfair treatment. The right to contest would be in addition to the ability to withdraw consent currently provided for in PIPEDA, or a "right to object" which functions in a similar manner. However, it is necessary to have both rights, as withdrawal of consent/right to object is an all-or-nothing decision, whereas contestation provides individuals with recourse even when they choose to continue to participate in the activity for which automated decision-making was employed.* [148]

The Canada ADM Directive does not establish an explicit right to contest automated decisions, but states that the Assistant Deputy Minister responsible for a program using an ADM system is responsible for:

> *6.4.1 Providing clients with any applicable recourse options that are available to them to challenge the administrative decision.*[149]

The Canada ADM Directive's commitment to "recourse options…to challenge administrative decisions" is commendable, but vague. It is not clear what "recourse options" entail, nor what remedies may be available. Moreover, while the Canada ADM Directive may acknowledge the ability to challenge administrative decisions, it does not actually *create* a legal right to do so. Rather, as Professor Scassa notes

> *While directives are important policy documents within the federal government, and while there are accountability frameworks to ensure compliance, the requirements to comply with directives are internal to government, as are the sanctions. Directives do not create actionable rights for individuals or organizations.*[150]

The Canada ADM Directive may be compared with a recent proposal in Washington State, House Bill 1655.[151] That proposed statute would create explicit appeal rights for individuals affected by AI government decisions and require agencies to be able to explain the bases for such decisions.[152] The Bill would also give any person injured by a "material violation" of the Act (which may include denial of any government benefit) a statutory right to institute proceedings against an agency and the right to seek injunctive relief, restoration of the government benefit in question, declaratory relief, or a "writ of mandate."[153]

The "right to contest" AI is gaining traction internationally. Most notably, section 3 of Article 22 of the European Union's General Data Protection Regulation establishes a right to contest certain decisions made by automated decision-making:

> *3. In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests,* ***at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.*** *[Emphasis added.]*[154]

As Kaminski and Urban note, establishing explicit AI contestability rights would be complicated. For example,

> *AI decisions are made at speed and at scale—features that in fact can be core justifications for using AI in the first place. To impose full judicial process on each AI decision would be to impose costs, both monetary and temporal, that might make the use of AI unwieldy. What, then, should be done to afford process that is not perfect, but instead is good enough?*[155]

At this point, the LCO is not recommending the adoption of an independent right to contest AI in Ontario. There are too many outstanding legal, technical and operational issues to address before we take that step. Moreover, the LCO believes there are more important law reform priorities in Ontario at this stage, including our recommended regulatory reforms and our recommendations in the areas of human rights and administrative law, discussed below. That said, there are important reasons to study the" right to contest AI" further and to learn from other jurisdictions.

The Relationship Between AI Regulation and AI Litigation

Discussions about AI contestability inevitably raise questions about the relationship between AI regulation, AI litigation and access to justice.

Systemic AI *regulation* and individual AI *litigation* are obviously different kinds of accountability tools: Systemic regulation establishes a legal accountability framework for all government AI decision-making within a jurisdiction. Litigation, on the other hand, is typically a means to challenge individual decisions made (or aided) by government AI systems.

There are many limitations to "regulation by litigation" as an accountability strategy. The LCO's Criminal AI Issue Paper discussed many of these limits at length, concluding that

> *Litigation has an important role in regulating AI and algorithms in the criminal justice system. Many issues will always be best addressed in open court with the benefit of an evidential record and high-quality and experienced counsel…*
>
> *Litigation, while obviously necessary to address specific cases, is insufficient to address the systemic statistical, technical, policy and legal issues that have been addressed in this report so far.*[156]

The LCO's analysis emphasized the enormous practical burden placed on individual defendants wishing to challenge bail or sentencing decisions based in part on algorithmic risk assessment tools:

> *Consider just some of the complex statistical, technical and policy issues that could (or should) have been litigated in Loomis or equivalent cases:*
>
> - *Is the historic data used to train the COMPAS tool biased, accurate, reliable and valid?*
> - *Are COMPAS risk factors and scores weighed and calculated appropriately?*
> - *Which communities bear the burden of statistical errors?*
> - *Are the confidence estimates for COMPAS predictions appropriate?*
> - *Are COMPAS predictions validated appropriately?*
> - *Does COMPAS use factors such as education or employment as impermissible statistical proxies for race or gender?*
>
> *Loomis was a comparatively simple, State-court criminal proceeding in which Loomis had already plead guilty. The COMPAS issue arose at sentencing. It is inconceivable that Loomis or any other criminal defendant (particularly one represented by a public defender/legal aid or self-represented) would be in a position to mount an effective challenge to the complex statistical, technical and legal issues raised by COMPAS.*[157]

The LCO subsequently concluded that legal accountability in criminal proceedings depends on both regulation of AI/ADM *systems* and a dedicated remedial regime that allows criminal defendants to effectively challenge *individual* AI/ADM decisions. The LCO warned that failure to adopt such measures could add significant barriers for low-income, Indigenous and racialized communities, thus compounding the over-representation of these communities in the criminal justice system.

Not surprisingly, the LCO has concluded that the lack of a dedicated regulatory and remedial regime in the civil and administrative law systems creates similar barriers. This is because the same issues arise in civil/administrative proceedings as in criminal proceedings: Absent systemic regulation of AI systems,

there is an enormous and perhaps insurmountable burden placed on litigants wishing to challenge individual AI-based government decisions. As a result, the LCO emphasizes that AI regulation and AI litigation are interdependent and that both are necessary to ensure legal accountability.

There is a further, perhaps underappreciated, relationship between systemic AI regulation and AI litigation: Proper regulation reduces litigation risk. This is because regulation significantly increases the likelihood that government AI systems will be designed, deployed and monitored with an eye to legal principles and requirements. This is yet another way in which thoughtful regulation and protecting individual rights complement and support each other.

The LCO does not want to overemphasis the salutary benefits of "regulation by litigation," even within the context of an appropriate regulatory framework. Even though there is an emerging community of practice in the US and Canada devoted to "litigating AI", it is widely acknowledged that AI litigation is likely to be very challenging.[158] For example, even extensive disclosure of an AI system may not be sufficient to ensure legal accountability because

> *…even perfect transparency into an algorithmic system—that is, unfettered access to its source code and data and the chance to observe its operation "in the wild"[91]—may not yield accountability in the sense of rendering decisions fully legible to data subjects or surfacing all of a system's flaws.[159]*

Moreover, irrespective of the level of disclosure, courts and tribunals assessing AI-based decisions may be "poorly situated to review the accuracy of the machine learning"[160] for several reasons, including

- Courts' and tribunals' potential lack of capacity to assess extraordinarily complex technical and statistical evidence; and,
- The difficulty of providing a remedy for systematic issues within the confines of an individual challenge.[161]

Finally, an AI system may change dynamically as the system is updated with new data or code. This situation may limit the precedential value of a court or tribunal decision if the AI system is significantly different at the time a decision is released.

This is not to say it is impossible that litigation will provide legal accountability for government AI systems in individual cases. The LCO strongly acknowledges the ability of courts and tribunals to address complex and novel technology issues in individual litigation. Litigating individual systems, however, will not provide the *systemic* legal oversight and accountability that the technology and rights at issue require. Nor will individual litigation be successful to address the access to justice challenges summarized below.

The Access to Justice Challenge

The LCO's mandate is to promote access to justice and evidence-based law reform. As a result, we must emphasize a significant drawback of "regulation by litigation": AI litigation places extraordinary legal and financial burdens on the individuals wishing to challenge government AI-based decisions.

Consider the access to justice implications of the following scenarios. The scenarios describe very common and important government functions: welfare determinations, bail hearings and child protection investigations:

- A government ministry has denied or cut off welfare benefits to a low-income Ontarian because a government AI risk assessment has predicted that she may be a high risk of fraud.

- A low-income, Black accused has been charged with assault. A provincial Crown is seeking to deny bail because a government AI risk assessment tool has predicted the accused is a high risk to either re-offend or to fail to appear at the bail hearing.
- An Indigenous child has been identified as being at high risk of suffering violence within his family. This assessment was made by a local child welfare agency relying, in part, on an AI system. The agency has begun an investigation into the family which may lead to further proceedings.

These scenarios are neither hypothetical nor improbable. Each one is based on a current example of an AI system widely used by governments in other jurisdictions that have been deployed for valid government objectives (fraud prevention, public safety and prevention of harm to a child).[162] In each case, the government AI system aggregates data from multiple data sources, embeds complex government policies into potentially millions of lines of code, and uses complex statistical modelling to weigh multiple variables to arrive at an individual risk score that is used to make, or inform, a government decision about an individual.

How can the individuals in these scenarios mount a meaningful challenge to the complex statistical, technical and legal issues raised in each example? A partial list of potentially relevant issues in each case might include:

- Is the data used to train the AI system biased, accurate, reliable or valid?
- Are the system's risk factors valid?
- Are the system's variables weighed and calculated appropriately?
- Does the system's code accurately reflect complex government policies?
- Is the system's operation or output understandable?
- Is the system accurate?
- Is the system testing and validation appropriate?

Challenging these complex issues in court or at a tribunal would be difficult under any circumstances, even assuming the challenging party had high quality, fully funded and technologically competent legal representation. These challenges would undoubtedly be significantly worse for under- or unrepresented litigants.

The legal and social implications of this analysis are worrisome. It is possible that only the best resourced and most sophisticated litigants will be able to challenge many AI-based government decisions. Absent positive initiatives, government AI decision-making may add significant new access to justice barriers to low-income, marginalized, Indigenous, and racialized communities, thus compounding the over-representation of these communities in Ontario's justice system.

Some key questions about government AI decision-making

The LCO's AI research has revealed consistent questions, issues, themes and principles affecting government AI systems. This table is a high-level summary of some of these key questions and issues.

Participation

- Are the public and stakeholders involved in the design, development, deployment, and oversight of the AI system? Is that participation meaningful?

Disclosure/Notice

- What types of government AI systems must be disclosed?
- What information will be disclosed?
- Will disclosure be both substantive and accessible?
- Will persons affected by an AI system be notified of its use?

Transparency

- Is the system sufficiently transparent to ensure the public understands its purpose, potential impact, operation, and evaluation?
- Is the system sufficiently transparent to meet or exceed legal procedural fairness standards and to ensure legal accountability?

Prohibitions, Risks and Harm Mitigation

- Should the system be prohibited due to systemic risks to legal rights and vulnerable populations?
- How is an AI system's risk assessed? And by who?
- Once a system's risk is identified, how will that risk be mitigated or eliminated?

Explainability and Reasons

- Can the operation and outcome of an AI system be explained to the public?
- Do AI decisions meet or exceed legal substantive fairness standards and ensure legal accountability?

Discrimination

- Does the AI system discriminate against vulnerable or human rights-protected populations?
- What steps will be taken to mitigate or eliminate the risk of discrimination?
- Is there regular testing and evaluation for bias?
- Does the system meet or exceed Charter or human rights legal standards?

Data Accuracy, Reliability, and Validity

- Is the data used to train an AI system accurate, reliable, and valid?

Human-In-The-Loop/Discretion

- Is there a "human-in-the-loop" reviewing a system's output and ensure oversight?
- Do human decision-makers have meaningful discretion to overturn or alter AI-based decisions or recommendation?
- How is automation bias addressed?

Evaluation and Monitoring

- How will system administrators and the public know the system is effective and legal?
- Is there an evaluation plan for the entire lifecycle of the system?

Human Rights

Most analysis of government AI systems to date has focused on AI decision-making in the criminal justice system or on specific technologies, such as facial recognition systems. This situation is beginning to change, however, in response to the growing use of these technologies to aid non-criminal government decision-making.

The remainder of this paper considers AI accountability in key areas of human rights, administrative law, privacy law and civil procedure. Government decisions in these areas range from minor, inconsequential decisions to major decisions affecting significant personal rights and interests. The Supreme Court of Canada has acknowledged that administrative decision-making is "one of the principal manifestations of state power in the lives of Canadians."[163]

Introduction

Human rights have been a transcendent issue in AI design, development, operation, and oversight across the world for many years.[164] The Toronto *Declaration Protecting the Right to Equality in Machine Learning* states the issue clearly:

> *As machine learning systems advance in capability and increase in use, we must examine the impact of this technology on human rights. We acknowledge the potential for machine learning and related systems to be used to promote human rights, but are increasingly concerned about the capability of such systems to facilitate intentional or inadvertent discrimination against certain individuals or groups of people. We must urgently address how these technologies will affect people and their rights. In a world of machine learning systems, who will bear accountability for harming human rights?*[165]

Although the human rights implications of AI are complicated and evolving, there is general agreement that human rights are crucial in the development of AI. There is also agreement that policymakers must move beyond recital of principles and focus operationalizing human rights compliance.

This section has two parts:

First, we identify the key human rights concerns, unanswered questions and legal gaps that arise when claimants want to challenge an AI system on human rights grounds.

Second, we identify many promising strategies for building human rights compliance and accountability into government AI systems. These strategies, when combined with the many strengths of Ontario's human rights system, suggest a promising path forward.

The LCO reiterates that the focus of this paper is on human rights and AI systems in the public sector. Human rights in private sector AI systems raise somewhat different issues and are not addressed here.

How Can AI Systems Affect Human Rights?

The use of AI systems can trigger a wide range of human rights issues. For example, surveillance technology can affect rights to privacy, freedom of expression or freedom of movement;[166] AI systems can restrict an. restrict an individual's autonomy, quality of life, or deny important benefits (such child welfare risk assessments or health care determinations). Finally, AI systems can increase barriers for people with disabilities if they are designed without considering accessibility.[167]

These issues are worthy of thoughtful analysis. In this paper, however, the LCO concentrates on government AI systems that may contravene the bias and discrimination provisions of the *Ontario Human Rights Code* and the *Charter*. The Partnership on AI, a leading American AI research institute, notes

> *A central concern with the rise of artificial intelligence (AI) systems is bias. Whether in the form of criminal "risk assessment" tools used by judges, facial recognition technology deployed by border patrol agents, or algorithmic decision tools in benefits adjudication by welfare officials, it is now well known that algorithms can encode historical bias and wreak serious harm on racial, gender, and other minority groups.*[168]

AI systems can be biased or discriminatory against individuals on the grounds of race, age, disability, sex, family structure or other protected grounds.[169] Bias in an AI system can also intersect across multiple grounds at once. Equally troubling, bias is often embedded, unexpected, undetected and coupled with the perception that machines are objective.

Discrimination can occur in the design of an AI system for many reasons, including the developer's personal bias or assumptions;[170] reliance on factors that correlate with bias (such as location data that correlates with race or employment data that correlates with gender);[171] or reliance on training data that is discriminatory.[172]

The most common human rights criticism of AI is the potential use of biased data. In these circumstances, because the training data or "inputs" used by an AI or algorithm (such as arrest, conviction, child welfare, education, employment or "fraud" data) may themselves be the result of biased practices, the results or outputs of an AI or algorithmic system may also be biased. In other words: "bias in, bias out."[173]

For many, the "bias in, bias out" argument is conclusive proof that AI or algorithmic tools should *never* be used in government decision-making. In this view, AI and ADM systems are often "a sophisticated form of racial profiling."[174] For others, AI or algorithmic tools are valuable because they have the potential to *reveal* systemic bias and discrimination. For example, many scholars and advocates believe that

> *With the appropriate requirements in place, algorithms create the potential for new forms of transparency and hence opportunities to detect discrimination that are otherwise unavailable. The specificity of algorithms also makes transparent tradeoffs among competing values. This implies algorithms are not only a threat to be regulated; with the right safeguards, they can be a potential positive force for equity.*[175]

The contrast between these perspectives – AI as perpetuating bias versus AI as revealing bias – runs through entire AI and human rights debate.

To its credit, the Government of Ontario has addressed this issue forthrightly: The government's 2019 discussion paper, *Promoting Trust and Confidence in Ontario's Data Economy*,[176] establishes promoting trust and confidence in the government's use of AI systems as the first of three pillars of the Ontario Data Strategy.[177] The paper notes bias and discrimination as a threat or risk associated with "data-driven practices" and links bias with "black box" concerns. The government's discussion paper provides two examples of algorithmic bias:

- An algorithm used to make an administrative decision assigns undue weight to a characteristic leading to discrimination against a certain group.
- A machine learning algorithm optimizes for numerically dominant groups in a training dataset, excluding data at the margins which represent a marginalized group.[178]

The "bias in, bias out" issue is the best-known AI bias issue, but not the only one. Discrimination and bias issues can also arise in questions regarding statistical "metrics of fairness", AI or algorithmic scoring, and automation bias, to name a few.[179] The effect of technology on bias can be subtle but significant. For example, AI may increase barriers for people with disabilities if systems are built without considering accessibility, or if they misinterpret a disability as cheating, an anomaly, or a red flag.[180]

This report will not repeat the LCO's earlier "data discrimination" analysis and recommendations.[181] Nor will the report discuss the many data science or technical initiatives designed to address bias in AI systems.[182] Rather, the LCO's focus here are the legal issues that arise from potentially discriminatory AI systems used by governments and public agencies in Ontario.

The legal system's ability to identify, litigate, and remedy discriminatory AI systems is crucial to establishing legally accountable AI in Ontario. Accordingly, this section discusses how Ontario laws of discrimination apply to AI, gaps in those laws, and steps policymakers and AI developers can take to mitigate discrimination.

Human Rights Law in Ontario

Ontarians' human rights are protected by the *Human Rights Code* (the *Code*),[183] the *Canadian Human Rights Act* [184] and the *Canadian Charter of Rights and Freedoms* (the *Charter*).[185] The *Code* focuses on equality and freedom from discrimination. The *Charter* is broader in scope. In addition to the right to equality and freedom from discrimination, it includes rights to liberty, to a fair trial, freedom of expression, freedom of mobility, among others.

Both the *Charter* and the *Code* protect individuals and groups from discrimination based on enumerated grounds such as disability, race, sex, age, and religion.[186] The *Code* applies to everyone in Ontario, including public and private entities, and focuses on social areas such as jobs, housing, services, unions and contracts. [187] The *Canadian Human Rights Act* protects against discrimination in federal jurisdiction, including discrimination in the federal government/programs and the federally-regulated private sector, including banking, railways and airlines. The *Charter* applies only to government action.

In Ontario, human rights are enforced through the Ontario Human Rights Tribunal and Ontario courts. The Ontario Human Rights Commission (OHRC) works to prevent discrimination and advance human rights through education, policy development, public inquiries and litigation.

Strengths in Ontario's Human Rights Framework

Ontario's existing human rights framework has many strengths for addressing AI-related discrimination issues:

First, the *Code* and *Charter* apply to decisions made by government that affect a person's rights. AI systems that make or aid government decision-making will therefore be subject to the *Charter* and Ontario's human rights legislation.

Second, courts have interpreted the *Code* and the *Charter* to address "substantive equality."[188] This suggests that under human rights law, government AI systems will be assessed in light of their impact on individuals and communities affected by a system[189] and whether a system in its entirety (including data, weighting, factors, outcome etc.) addresses people's circumstances, historic experiences, and ongoing barriers.[190]

Third, it is not necessary for discrimination to be intentional to contravene Ontario's human rights laws. Human rights law in Ontario protects against indirect discrimination. In other words, where a system, law or policy is equal on its face, but in practice has a discriminatory effect, the system can violate human rights laws.[191] This means governments are not shielded from human rights liability because an AI system's developers or administrators did not know the system was discriminatory.

Fourth, it is not necessary for discrimination to affect all members of a protected group in the same way. The Supreme Court of Canada has said that "The fact that discrimination is only partial does not convert it to non-discrimination."[192] For example, a plan which denies benefits during pregnancy is discriminatory on the basis of sex, even if it only applies to "women who are pregnant" and not all women.[193]

Fifth, the constitutional status of the *Charter* and quasi-constitutional status of the *OHRC*[194] mean they take precedence over other laws and regulations if there is a conflict. Human rights laws are to be given a large and liberal interpretation to ensure they are upheld and exceptions to human rights protections are to be "narrowly construed."[195]

Sixth, the Ontario Human Rights Commission can draft policy guidelines[196] or be added as an intervenor/party to a test case where new law is likely to be made.[197] The Ontario Human Rights Tribunal has sweeping inquiry powers[198] and can go beyond restitution for an individual and order respondents to correct systemic issues.[199]

Finally, Ontario has a broad and capable network of NGOs, community legal clinics, lawyers, judges, tribunal members and many others dedicated to protecting human rights across the province.

Challenges and Unresolved Legal Issues in Ontario's Human Rights Framework

Notwithstanding these strengths, the LCO believes there are many challenges and unresolved legal issues in Ontario's human rights framework governing AI systems. Many of these challenges/questions are systemic; others arise in the context of human rights litigation.

Before beginning this analysis, it is important to set out the basic legal elements necessary to prove discrimination in Ontario. To establish legal discrimination, a claimant must show:

- They have a characteristic protected by the *Code* or *Charter* (e.g., race, age, sex);
- They experienced adverse treatment or impact; and
- The protected characteristic was a factor in adverse treatment or impact.[200]

When litigating an AI-based discrimination claim, an applicant must make out a *prima facie* case on a "balance of probabilities." Once a *prima facie* case is established, the burden shifts to the respondent to justify the conduct within the framework of the exemptions available under the *Code*,[201] or in the context of a *Charter* claim, whether the discrimination can "be justified in a free and democratic society."[202]

Using this framework to protect human rights in government AI raises many new issues and is likely to be challenging. The following discussion addresses several relevant issues, including

- The difficulty of identifying and proving systemic discrimination in AI cases
- Unresolved and novel data and statistical issues
- Unresolved issues regarding accommodation in AI systems
- Access to justice challenges, especially for self-represented litigants

The Supreme Court has described discrimination as a "distinction" relating to personal characteristics of the individual or group, which has the effect of imposing burdens, obligations, or disadvantages on such individual or group not imposed upon others, or which withholds or limits access to opportunities, benefits, and advantages available to other members of society.[203]

The Ontario Human Rights Commission defines systemic discrimination as including the following three elements: (i) patterns of behavior, policies or practices; (ii) part of the social or administrative structures of an organization; and (iii) position of relative disadvantage created for persons identified by the *Code*."[204]

Based on these definitions, discrimination in government AI systems is likely to fulfill the OHRC's definition of systemic discrimination for the following reasons:[205] AI systems are designed to detect patterns and categorize information based on these patterns; government use of an AI system would be a policy or a practice and its consistent application would make it part of the administrative structure of the organization; and AI systems may categorize on the basis of prohibited grounds, even though they are not designed to.

Proving Systemic Discrimination

Even in normal circumstances, systemic discrimination is difficult to establish due to several factors:

- **Systemic discrimination can take many forms and is potentially undetectable to those unaffected.** The Supreme Court and Human Rights Commissions recognize that "disadvantage arises from the way in which society treats particular individuals, rather than from any characteristic inherent in those individuals"[206] and that "discrimination may be based as much on other people's perceptions, myths and stereotypes, as on the existence of any actual [characteristic of] the individual."[207]
- **Human rights and equality law is complex.** Human rights and equality law is often confounding, even to constitutional law scholars and human rights experts.[208] Judicial analysis of discrimination is contextual, purposive, flexible, varied and broad.[209]
- **Evidential standards can be challenging.** Establishing evidence to prove systemic discrimination is challenging for many reasons. As a starting point, discrimination is often not overt and can be imperceptible to those unaffected. Discrimination that is inadvertent – action that is neutral on its face, but indirectly burdens one population more so than others – can

be very difficult to prove. Second, information and knowledge about a "system" may be unavailable to a plaintiff. Assembling evidence about a system is resource intensive and often requires considerable data and research.

Courts and tribunals recognize evidentiary challenges faced by plaintiffs in discrimination cases.[210] The evidentiary bar to establish a prima facie case of discrimination is unclear. Some courts describe the threshold as "extremely low"[211] or "not high".[212] Others, however, suggest the bar for proving systemic discrimination is quite high.[213]

Requirements for evidence in discrimination cases has evolved. Most recently, in *Fraser* the Supreme Court sought to clarify the evidence required for a party to prove indirect discrimination. Justice Abella stated that courts will look at:

(i) Evidence about the situation of the claimant group, including the physical, social, cultural or other barriers.[214] Issues which predominantly impact certain populations may be under-documented, and courts may have to rely more heavily on claimant's own evidence or evidence from other members of their group, rather than on government reports, academic studies, or expert testimony.[215]

(ii) Evidence about the law, including evidence that may provide proof that members of a protected group are disproportionately impacted.[216] In this manner, courts will look to evidence about the results of the impugned law or government action, including statistics. The goal of statistical evidence is to establish a "disparate pattern of exclusion or harm that is statistically significant and not simply the result of chance."[217]

The weight given to statistics depends primarily on quality and methodology.[218] There is no "universal measure" for what level of statistical disparity is necessary to demonstrate disproportionate impact.[219] The statistics must show a pattern that is significant and not just "the result of chance".[220]

Justice Abella clarifies that ideally claimants will produce both kinds of evidence, but there may be some circumstances where both types are not required.[221]

As discussed above, the framework for making a prima facie case of discrimination requires the applicant to show that the protected characteristic was a "factor" in the adverse treatment,[222] or that the differential treatment was "based" on a protected ground.[223] In *Fraser*, Justice Abella clarified that as long as a claimant demonstrates disproportionate impact on a protected group, they need not also prove that there is a causal link between the two.[224] There does, however, need to be some nexus between the protected ground and adverse impact.[225] Even without having to show a "casual link", this standard can still prove to be a significant evidentiary challenge in systemic discrimination cases.[226]

AI and Systemic Discrimination

Proving systemic discrimination in any case is difficult. Establishing systemic discrimination in AI cases is likely to be harder still. Professor Teresa Scassa states that

> *…in order to establish discrimination [in an AI system], it will be necessary either to demonstrate discriminatory impacts or effects, or to show how the algorithm itself and/or the data used to develop it incorporate biases or discriminatory assumptions.*[227]

This will be a difficult burden to meet. AI adds complexity to existing systemic discrimination issues and raises several novel and difficult questions, including:

- The unique features of government AI decision-making (speed, scale, the "black box", complexity, embedding decisions into code, data issues, etc.) mean that identifying (or suspecting) systemic discrimination may be difficult.
- The opacity and confidentiality of an AI system/data may mean that a plaintiff cannot access relevant evidence.
- The statistics needed to show that an AI system has a disproportionate impact on certain populations will require records which may not be available to a plaintiff or maintained at all.
- Technical information about an AI system may be unrelatable or uninterpretable to plaintiffs, counsel and triers-of-fact. For some AI systems, outcome data may not be explainable even to the persons or organizations that designed, built and operated the system.
- Systems are constantly changing and being updated. As a result, the data or systems being challenged may be out of date at the time of litigation.
- It may be difficult to establish a link between the AI system and government decisions affecting individuals, especially if the system is opaque or unexplainable.[228]

The sum of these challenges is that meeting the systemic discrimination evidential burden in government AI cases will very likely require considerable resources, extensive disclosure and technical expertise. This emphasizes the need for regulation as some of these challenges can be overcome through regulation. For example, if operators of AI systems are required to test, audit and validate their systems throughout its life cycle and disclose the results, biased systems may be uncovered prior to deployment. In some ways, AI may make it easier to show statistical disparity in systemic decision-making.

Unresolved and Novel Statistical Issues

In addition to the statistical issues discussed above, AI controversies in the United States have highlighted little known but important questions concerning the "metrics of fairness."[229] Canadian courts and tribunals may need to address similar issues as part of their human rights analysis of government AI systems.

The highest profile illustration of the "metrics of fairness" issue is the COMPAS controversy in the United States. COMPAS was an algorithmic tool used by many American jurisdictions to predict recidivism in U.S. bail and sentencing proceedings.[230] In 2016, ProPublica, an American investigative journalism organization, published an article titled "Machine Bias" that summarized ProPublica's research on the use of the COMPAS risk assessment tool to inform criminal sentencing in Broward County, Florida.[231] The ProPublica article was damning:

> *…We also turned up significant racial disparities…In forecasting who would re-offend, the algorithm made mistakes with black and white defendants at roughly the same rate but in very different ways.*
>
> *The formula was particularly likely to falsely flag black defendants as future criminals, wrongly labeling them this way at almost twice the rate as white defendants.*
>
> *White defendants were mislabeled as low risk more often than black defendants.*[232]

ProPublica concluded by stating that COMPAS was "biased against blacks" because Black defendants were over-classified as risky.[233] The ProPublica analysis was vigorously debated on both methodological and policy grounds.[234] Nevertheless, the ProPublica article "supercharged" the emerging debate on risk assessments and racial bias.[235]

The COMPAS controversy demonstrated how different measures of statistical fairness are crucial in determining whether an algorithm or AI system should be considered discriminatory or race-neutral. The controversy also demonstrated that the burden for an algorithm's or AI system's statistical errors may not be shared equally: a statistical measure that over-classifies racialized accused as risky may effectively replicate (or worsen) existing patterns of racial disparity.

The American discussion about the relationship between statistical measures of fairness and legal principles is just beginning. This debate is unlikely to be resolved soon or conclusively, as there appear to be at least *six* different statistical measures that can be used to evaluate some algorithmic tools.[236]

At this point there is no clear framework for how to evaluate measures of equality in statistics or how to evaluate if an AI system or algorithm is biased or not. According to Professor Sandra Mayson, "[American] law provides no useful guidance about which [statistical measure] to prioritize."[237] Professor Aziz Huq goes further, stating that American constitutional law "provides no credible guidance "as to how to assess racial equity in algorithmic risk assessments.[238]

Policymakers, courts, tribunals, lawyers, human rights experts and technologists in Ontario should monitor the American debate, but be cautious about its implications in Canada.

One of the challenges with trying to measure statistical "fairness" is the inherent complexity of measuring a concept that is multi-faceted and subjective. In some respects, the lack of a clear metric to measure fairness in AI and algorithms accords with current Canadian human rights jurisprudence. The Supreme Court of Canada has stated that there is no "universal measure" for what level of statistical disparity is necessary to demonstrate disproportionate impact.[239] Statistics must show a pattern that is significant and not just "the result of chance."[240] The weight given to statistics also depends on quality and methodology.[241]

Although there may be benefits to creation of a standard for how to measure discrimination in AI, there are pitfalls as well. The Supreme Court has clearly stated that "…the Court should not…craft rigid rules" to determine a "universal measure for what level of statistical disparity is necessary to demonstrate that there is a disproportionate impact."[242] Justice Abella's reasoning in *Fraser* can apply also in the context of AI systemic discrimination claims: Data can produce numbers which seem on their face to be clear and neutral, but do not address nuances in the human experience. Data may or may not include information about protected characteristics.

The LCO does not believe that explicit or numerical "standards" to assess bias in AI should be created at this time. However, discussions about monitoring and auditing AI for bias should include analysis of how to reconcile the complexities of statistical measures of fairness in AI systems with legal analysis of disproportionate impact.

Accommodation in Government AI Systems

Applying a human rights framework to government AI systems will require courts and tribunals to consider accommodation issues.

At present, courts and tribunals apply a three-part test to assess whether a discriminatory action can be justified:[243]

- The discriminatory rule or action must have a purpose or goal that is rationally connected to the function being performed.
- A reasonable and bona fide justification for discrimination requires "an honest and good faith belief in the necessity of the rule. [244]
- The discriminatory action must be reasonably necessary. Under this section the defendant must show that it is impossible to accommodate the claimant without undue hardship. [245]

Tribunals and courts will have to determine how to apply this test in the context of litigating an AI system. In so doing, tribunals and courts will be challenged to answer several complex and vexing questions, including:

- Can a government AI system that is shown to discriminate on enumerated grounds ever be necessary or permitted?
- How can governments demonstrate that *not* using an AI system would cause undue hardship?
- Can reasonable accommodation requirements be met with an AI system opt-out procedure?

In the case of a *Charter* violation, governments will have the opportunity to argue that the breach is "justified in a free and democratic society."[246] In these circumstances, courts will need to balance the severity of the discrimination against the program's statutory objectives.[247]

Where an AI system is operating pursuant to an administrative body's enabling statute, AI decisions or recommendations are likely to be described as an exercise of administrative discretion. A *Charter* challenge to administrative decision maker's exercise of discretion are assessed through a modified framework. [248] Judicial review of these decisions focusses on proportionality.[249]

Some scholars believe the proportionality test makes it too easy for government to defend *Charter* breaches.[250] The proportionality test is also criticized for lacking clarity[251] or because it is applied unpredictably.[252] The paucity of cases challenging administrative discretion on *Charter* grounds could also suggest that "there are a great number of cases that engage in a 'conspiracy of silence' where counsel – likely making a calculated, cost-saving decision based on experience and expertise – elects not to raise a *Charter* argument."[253]

It is worth noting that statutes that *may* be interpreted in a manner that violates the *Charter* are constitutionally valid.[254] Courts can select to interpret an overbroad or ambiguous statute as constitutional.[255] Scholars have argued that this constitutional presumption "can complicate *Charter* litigation for the litigant"[256] because litigants in these circumstances can only access personal remedies.

This analysis suggests that litigants may face an uphill battle obtaining systemic remedies against AI systems. For example, if a party is seeking to have an AI-related provision declared unconstitutional, it will likely be insufficient to simply show that a government AI system could breach the *Charter*.

Other AI and Human Rights Issues

Finally, the LCO wants to highlight two further AI and human rights issues that may arise in AI cases: intersectionality and AI used as an investigatory or research tool.

Intersectionality

People who identify with more than one enumerated ground face a complexity of prejudices that can be unique to their circumstances. Although there is currently no explicit recognition of intersectionality in the Code[257] individuals can assert claims in respect of each protected ground, and the Tribunal accepts the vulnerability of intersecting Code grounds.[258] The Tribunal recognizes the difficulty in arguing and proving this type of discrimination.[259]

The challenge of intersectionality in human rights law is likely to be amplified with AI systems in two ways. First, it may be impossible to decipher or understand how the output from certain AI systems is prejudicial or discriminatory to individuals who cross multiple enumerated grounds. Second, disadvantaged and marginalized members of society are likely to be more heavily impacted by AI systems. For example, some individuals or communities could be affected by AI systems used in policing, child welfare and social benefits. The impact AI will have on these individuals and communities could be very significant.

AI as an Investigatory or Research Tool

There is grey area regarding how human rights law could apply to government AI systems that do not directly affect an individual, including AI used to identify or prioritize investigations, to inform research or policy decisions, or for internal government management. For example, governments are increasingly using AI systems to prioritize investigations for regulatory violations, to identify tax records to audit, or to identify potential securities violations. The U.S. Securities and Exchange Commission is a notable leader in this field.[260]

These applications potentially raise human rights concerns, even though their impact on individuals or communities may not be readily apparent. Government use of AI in these contexts should be used thoughtfully so not to embed undetected or unknown discrimination, especially when there is little transparency or redress for individuals who may be affected.

Strategies to Promote Human Rights in Public AI Systems

AI bias raises the question of how the developers and administrators of government AI systems can design and maintain systems to minimize or avoid discrimination.

In January 2022, the Ontario government released *Beta Principles for the ethical use of AI and data enhanced technologies in Ontario*.[261] The Beta Principles explicitly state that "Data enhanced technologies should be designed and operated in a way throughout their life cycle that respects the rule of law, human rights, civil liberties, and democratic values. These include dignity, autonomy, privacy, data protection, non-discrimination, equality, and fairness."

These are easy recommendations to make. They are a challenge to operationalize.

In late 2021 and early 2022, the LCO ran a series of workshops on AI, legal issues and regulatory considerations with government policymakers, operational staff and AI developers. Bias, discrimination and human rights law was of considerable interest to all participants. There was a consensus about the importance of designing and operating AI systems free from discrimination, but many uncertainties about how to achieve this objective. The sessions considered many of the issues highlighted in this section: What human rights laws apply? What do human rights laws mean in practice? How can these legal principles be incorporated into the design, development, deployment and oversight of government AI systems?

This section summarizes an early set of initiatives and strategies to build "human rights by design" into government AI systems and to ensure human rights accountability for systems.

It is important to note that these questions are being addressed by experts around the world.[262] The LCO's focus is on Ontario, however, and steps developers and policymakers in this province can take to ensure human rights compliance and legal accountability.

Potential strategies include:

Public Participation
Some of the notable proposals suggested for achieving human rights accountability and compliance include ensuring tools are designed with input from members of the local community, including, but not limited to:

- Independent data scientists.
- Community groups focused on racial justice.
- Establishing community advisory boards.
- Regular training on tools.

Constitutional or Human Rights Provisions
Several jurisdictions have adopted or considered explicit legislative commitments to ensure AI and ADM systems are compliant with constitutional law or anti-discrimination statutes. These provisions may include legislative findings, preambles or explicit provisions stating that an AI or ADM system must comply with constitutional principles or anti-discrimination legislation. For example, Washington State House Bill 1655 includes provisions that

> *A public agency may not develop, procure, or use an automated decision system that discriminates against an individual, or treats an individual less favorably than another, in whole or in part, on the basis of one or more [enumerated] factors…*

Data Disclosure Requirements
Many of the proposals and options promoting greater disclosure of AI and ADM systems are designed to reduce bias and discrimination. Regulatory proposals that explicitly require disclosure of data are directly linked to concerns about the potential impact of racialized data in AI and ADM systems.

Accordingly, many of the proposals for addressing bias and discrimination require comprehensive disclosure of the data and variables used to train, calibrate, and operate AI and ADM systems. Data disclosure requirements could include:

- Training data
- Description of design and testing policies and criteria
- List of factors that tools use and how they are weighted
- Thresholds and data used to determine labels for scoring
- Outcome data used to validate tools
- Definitions of what the instrument forecasts and for what time period
- Evaluation and validation criteria and results

Human Rights or Bias Impact Assessments/Best Practice Guidelines
In addition to the proposals discussed above, there are many emerging best practices for addressing bias and discrimination in AI and ADM systems.

For example, the European Commission's High Level Expert Group's Assessment List includes a questionnaire.[263] Key excerpts include:

> ***Avoidance of Unfair Bias***
> - *Did you establish a strategy or a set of procedures to avoid creating or reinforcing unfair bias in the AI system, both regarding the use of input data as well as for the algorithm design?*
> - *Did you consider diversity and representativeness of end-users and/or subjects in the data?*
> - *Did you test for specific target groups or problematic use cases?*

- *– Did you research and use publicly available technical tools, that are state-of-the-art, to improve your understanding of the data, model and performance?*
- *– Did you assess and put in place processes to test and monitor for potential biases during the entire lifecycle of the AI system (e.g., biases due to possible limitations stemming from the composition of the used data sets (lack of diversity, non-representativeness)?*
- *– Where relevant, did you consider diversity and representativeness of end-users and or subjects in the data?*
- *Did you put in place educational and awareness initiatives to help AI designers and AI developers be more aware of the possible bias they can inject in designing and developing the AI system?*
- *Did you ensure a mechanism that allows for the flagging of issues related to bias, discrimination or poor performance of the AI system?*
- *– Did you establish clear steps and ways of communicating on how and to whom such issues can be raised?*
- *– Did you identify the subjects that could potentially be (in)directly affected by the AI system, in addition to the (end-)users and/or subjects?*
- *Is your definition of fairness commonly used and implemented in any phase of the process of setting up the AI system?*
- *– Did you consider other definitions of fairness before choosing this one?*
- *– Did you consult with the impacted communities about the correct definition of fairness, i.e., representatives of elderly persons or persons with disabilities?*
- *– Did you ensure a quantitative analysis or metrics to measure and test the applied definition of fairness?*
- *– Did you establish mechanisms to ensure fairness in your AI system?*[264]

Another example is from the Brookings Institution, who developed a template "Bias Impact Statement"[265] with the following questions:

What will the automated decision do?

Who is the audience for the algorithm and who will be most affected by it?

Do we have training data to make the correct predictions about the decision?

Is the training data sufficiently diverse and reliable? What is the data lifecycle of the algorithm?

Which groups are we worried about when it comes to training data errors, disparate treatment, and impact?

How will potential bias be detected?

How and when will the algorithm be tested? Who will be the targets for testing?

What will be the threshold for measuring and correcting for bias in the algorithm, especially as it relates to protected groups?

> **What are the operator incentives?**
>
> What will we gain in the development of the algorithm?
>
> What are the potential bad outcomes and how will we know?
>
> How open (e.g., in code or intent) will we make the design process of the algorithm to internal partners, clients, and customers?
>
> What intervention will be taken if we predict that there might be bad outcomes associated with the development or deployment of the algorithm?
>
> **How are other stakeholders being engaged?**
>
> What's the feedback loop for the algorithm for developers, internal partners and customers?
>
> Is there a role for civil society organizations in the design of the algorithm?
>
> **Has diversity been considered in the design and execution?**
>
> Will the algorithm have implications for cultural groups and play out differently in cultural contexts?
>
> Is the design team representative enough to capture these nuances and predict the application of the algorithm within different cultural contexts? If not, what steps are being taken to make these scenarios more salient and understandable to designers?
>
> Given the algorithm's purpose, is the training data sufficiently diverse?
>
> Are there statutory guardrails that companies should be reviewing to ensure that the algorithm is both legal and ethical?

Explanation Requirements

The Government of Canada's Guideline on Service and Digital contain important but perhaps underappreciated provisions linking requirements for explanations and interpretable models with bias reduction. For example, the federal Directive on Service and Digital states:

> *Having an easily interpretable model can also greatly simplify testing and monitoring of [an ADM] system, including assessing bias.*[266]

Research, Testing and Evaluation Requirements

Another important initiative to reduce the potential for bias and discrimination is a requirement for regular research, testing and evaluations of AI and ADM systems. Measures that have been suggested include:

- Comparing outcomes for different groups
- Testing for equality of outcome rates
- Identifying different definitions of fairness

The Canada ADM Directive requires that appropriate testing processes, designed to ensure that training data excludes unintended biases "and other factors that may unfairly impact the outcomes" be developed before the system goes into production. For example, data used by the system must be "routinely tested to ensure that it is still relevant, accurate, and up-to-date."[267] The Directive also requires peer review of all systems classified at Level II or higher. Peer review for Level IV

systems is notably robust, requiring the review of at least two qualified experts from listed organizations, including the National Research Council of Canada, relevant non-governmental organizations, or a contracted third-party vendor with a relevant specialization.

Concluding Thoughts on AI and Human Rights

Government AI systems must be human rights compliant. However, achieving human rights compliance will be difficult unless several important issues are addressed.

AI systems can be biased or discriminatory against individuals on many grounds including race, age, disability, sex, and family structure. Bias in AI systems can also intersect across multiple grounds at once. Discrimination can occur in the design of an AI system because of the developer's assumptions. It can also occur through use of data with biases often hidden or embedded into the system.

In extensive LCO consultations with government officials, it was apparent that AI developers, administrators and policymakers are committed to developing and operating government AI systems that comply with human rights law. The provincial government has publicly affirmed this goal. No one wants to knowingly create or deploy a biased AI system.[268]

Notwithstanding its many strengths, the LCO has concluded that the current human rights framework in Ontario is not sufficient to protect human rights where AI systems are relied on for government decision making. In the long run, human rights compliance will depend on how policymakers, courts and tribunals address significant evidential challenges inherent in "black box" government AI systems. These challenges can be addressed, in part, by systemic and significant disclosure and transparency of government AI systems.

Human rights compliance will also depend on thoughtful answers to several equally important legal, technical and practical issues, including:

- Data standards
- Evidential standards
- Guidelines or metrics to measure bias and discrimination in AI systems
- Bias testing or auditing requirements
- Determining reasonable accommodations in AI systems
- Remedies
- Access to justice challenges

Finally, human rights compliance of government AI systems will depend on addressing two further issues: First, can AI systems be used to reveal or address systemic discrimination? Second, are there AI "no-go" zones where a government AI system's potential risk to human rights is so significant that governments should prohibit the use AI in that area?

As mentioned earlier, the LCO believes there is an urgency to addressing these issues proactively. The growing use of government AI systems will likely result in more public interest in human rights issues in Ontario. There is also likely to be more human rights litigation challenging government AI systems in the coming years. The increased use of these systems, when combined with the novel legal issues they present, make them a target for systemic discrimination challenges.

Fortunately, there are many promising practices and law reform measures to begin the process of ensuring human rights compliance in government AI systems. An important early initiative could be to develop a made-in-Ontario AI Human Rights Impact Assessment to assist developers, policymakers, decision-makers and the public assess the human rights compliance of a government AI system.

Further steps include:

- Requiring human rights experts and communities to be involved in the design, development and operationalization of government AI systems
- Requiring human rights experts and communities to be meaningfully engaged throughout the lifecycle of a government AI system
- Requiring bias testing or auditing of AI systems

The LCO stresses that these recommendations supplement, and in no way replace, our earlier recommendations regarding an appropriate regulatory framework for government AI systems in Ontario. Those recommendations include several reforms (such as mandatory disclosure of AI systems, risk assessments, etc.) that provide the foundation for the human rights-specific recommendations identified here.

Absent appropriate policy guidance on human rights AI issues, provincial ministries, agencies, tribunals and/or courts will likely need to address complex legal and technical issues on a case-by-case basis, which may result in poorer public services, inconsistent decision-making, diminished rights protection, delays, added costs and unnecessary litigation.

Finally, it is important to note the role of the Ontario Human Rights Commission and Human Rights Tribunal in AI discussions. The OHRC has the authority to draft policy guidelines and to add themselves as an intervenor or a party to a test case where new law is likely to be made. The Human Rights Tribunal can go beyond restitution for a single individual and order respondents to correct systemic issues. In these circumstances, the Ontario Human Rights Commission and Human Rights Tribunal will play an important role in determining if, or how, systemic discrimination is addressed in government AI systems.

Fortunately, the OHRC has already begun to consider these issues. On December 10, 2021, the LCO, the Ontario Human Rights Commission and the Canadian Human Rights Commission announced a joint research and policy initiative to examine human rights issues in the development, use and governance of artificial intelligence and algorithms in Canada and specifically in Ontario. Over the coming years, our organizations will collaborate to research and co-develop resources to help identify and consider discrimination and other human rights issues in the use of artificial intelligence. We will provide public updates on the progress of this initiative.

In order to ensure government AI systems comply with human rights requirements, the LCO makes the following recommendation:

Recommendation 12

The provincial government, Ontario Human Rights Commission, technologists, human rights experts, and community members work together to develop a provincial human rights strategy for Government of Ontario AI systems.

This strategy should address the following issues:

- **A made-in-Ontario AI Human Rights Impact Assessment**
- **Data standards**
- **Evidential standards in government AI systemic discrimination cases**
- **Guidelines or metrics to measure bias and discrimination in government AI systems**
- **Bias testing or auditing requirements**
- **Determining reasonable accommodation in government AI systems**
- **Reviewing remedy provisions in the Ontario Human Rights Code for sufficiency to address potential harms of government AI systems**
- **Access to justice challenges**

A key element of this strategy should be to develop guidance for policymakers to determine what AI systems or applications should be prohibited on human rights grounds.

This strategy should require:

- **That human rights experts and communities be involved in the design, development, and operationalization of government AI systems**
- **That human rights experts and communities be meaningfully engaged throughout the lifecycle of a government AI system**

Administrative Law

Government decisions about individuals and organizations are significant and ubiquitous. Administrative law is the area of law that holds governments accountable for these decisions.

In simple terms, there are three key stages in challenging an administrative decision made by a government, public agency, or regulatory tribunal:

- The first stage is to assess whether there are any objections
- The second stage is a review of the decision on procedural and/or substantive grounds
- The third stage considers remedies

The sections that follow largely address the second issue: procedural and/or substantive review of government decisions.

As will be seen below, there are many unanswered questions about how administrative law will respond to government AI decision-making.

Procedural Fairness

Government decisions that affect an individual's rights, interests or privileges require procedural fairness.[269] Many early AI systems have failed judicial scrutiny on due process grounds. For example, the Arkansas and Houston systems discussed earlier were both challenged on due process grounds.

The principles of procedural fairness (fairness, transparency, accountability) and the safeguards created to achieve these principles (notice, participation, impartiality, reasons, rights of appeal) have largely been established by the Supreme Court of Canada in a series of seminal decisions.[270] These principles are the same for all administrative decisions, but how they are applied differs depending on the regulatory body, their governing legislation, and the decision in question.

A decision by government that affects the rights, privileges or interests of an individual triggers a duty of fairness.[271] The duty of fairness has been held to translate into an entitlement to fair procedure, which is founded on open decision-making and transparency.[272] Fair procedure is achieved through:

- Notice of the case one needs to meet
- Participation in the form of an oral hearing and/or written submissions
- A decision by an open-minded decisionmaker who considers the factors they are supposed to consider
- An explanation or reasons for the decision.

Courts have held that the requirements of procedural fairness are "eminently variable" and heavily context specific.[273] For example, the Supreme Court has held that an oral hearing is not always necessary to "ensure a fair hearing and consideration of the issues involved. The flexible nature of the duty of fairness recognizes that meaningful participation can occur in different ways in different situations."[274]

Similarly, although courts often emphasize the "usefulness of reasons in ensuring fair and transparent decision-making,"[275] the duty of fairness does not require that reasons be provided for all administrative decisions.[276] These safeguards are applied fluidly depending on the nature and significance of the decision, the expectations of the parties involved, and the specific statutory scheme and procedure of the administrative body at issue.[277]

AI introduces several layers of complication to these already difficult issues.

AI Systems and Procedural Fairness

One of the most important strategies for addressing concerns about AI accountability, transparency and fairness is to frontload AI systems with procedural fairness safeguards. Fortunately, the Government of Canada has taken important first steps in this regard. The Canadian ADM Directive and its companion, the Algorithmic Impact Assessment ("AIA"), are leading examples of tools and strategies that incorporate procedural fairness protections into the design and operation of automated government decision-making.[278]

The Canada ADM Directive is analyzed at length in the LCO's *Regulating AI* Issue Paper.[279] The Canada ADM Directive was crafted upon the principles of procedural fairness. It applies to federal agencies subject to the Treasury Board who intend to introduce any automated decision-making tool that could impact the rights, privileges or obligations of Canadians.

Professor Teresa Scassa analyzes the Directive and AIA in a recent paper where she considers the differences between the existing administrative law principles and the Directive.[280] Highlights of her analysis include:

Procedural Fairness is a Contextual Analysis

Professor Scassa notes that the Directive acknowledges the contextual nature of administrative law by adopting a sliding scale of "impact" categories, ranging from Category I (which includes algorithms that will have "little or no impact") to Category IV (which includes algorithms that will have a "very high impacts"). An algorithm's impact is determined by the Algorithmic Impact Assessment (AIA) tool, which asks the system's developers and sponsors to address a comprehensive services of questions about the proposed system.[281] The Directive also includes an Appendix which specifies a sliding scale of obligations or harm mitigation strategies that accompany each risk level.[282]

Notice

The Directive requires that notice be provided where a decision "will be undertaken in whole or in part" by an ADM system. For the lowest level of impact (Category I), no individual notice is required. For high impact (Category III) or very high impact (Category IV) decisions, the notice requirements are more onerous. The Directive specifies that notice requirements include plain language information

posted on a "relevant website" setting out how the ADM system works, information about how the system supports the administrative decision, the results of reviews or audits of the system, and either a description or a link to the anonymized training data used in developing the system.[283]

Participation

The Directive does not explicitly address participation or the right to a hearing. Rather, Professor Scassa comments that the Directive appears to adopt a right to human involvement.[284] According to the Directive, a "human-in-the-loop" is not required for decisions with low to moderate impact. [285] Conversely, decisions with high or very high impact must include human intervention during the decision-making process and the final decision must be made by a human.[286]

The Decision-Maker

Administrative law requires that the person making the decision has authority to do so.[287] Professor Scassa says therefore

> *a first question in the ADM context is … whether, if enabling legislation designates a human decision-maker or a decision-making body constituted of humans, the use of an ADS to decide matters – in whole or in part – is an improper delegation.* [288]

Some scholars have concluded that an AI system cannot replace a human decision-maker unless the enabling statute specifically provides for it.[289]

Impartiality

Impartiality in traditional administrative decision-making focuses on the individual making the decision. Administrative tribunals are not required to be impartial, but they must have an open-mind before hearing a party's submissions.[290] Professor Scassa points out that freedom from bias in this context is "about humans and their ability to be swayed or influenced,"[291] naturally raising questions about how to evaluate/address bias when machines are involved.[292]

The Directive does not account for potential issues of bias when there is a human-in-the-loop, such as questions about automation bias. However, the Directive does address bias from the perspective of systemic discrimination. The Directive requires peer-review of automated decision-making systems that have a moderate to very high potential impact. The extent of peer review depends upon the level of potential impact.[293]

Right to Reasons

The Directive requires that decisions be provided with a "meaningful explanation."[294] High to very high impact decisions must always provide an explanation of how, and why, the decision was made directly to the individual. Decisions with moderate impact are required to provide explanation on request. The lowest impact decisions can provide a "meaningful explanation" by posting to a general FAQ website.[295]

Importantly, Professor Scassa points out that an "explanation" is not the same as "reasons." The former "relates to the decision-making process (or algorithm)", while the latter "relates more to the motivations of the decision maker."[296]

Does the Directive Comply with Procedural Fairness?

It is important to ask whether or how the Directive complies with current principles of Canadian administrative law.

In many ways, the Directive *raises* the standard of procedural fairness:

- The Directive's explicit commitment to procedure fairness in algorithmic decision-making is both notable and unique in AI/ADM governance tools to date.
- The AIA tool is a sophisticated effort to analyze the impact of an ADM-aided decision, giving structure and systemic analysis to what might otherwise be an ad hoc, localized process lacking transparency and rigour.
- The requirements of notice and explanation could potentially provide greater transparency than what is currently required under common law.[297]
- The Directive includes explicit (and contextual) requirements and obligations that must be addressed depending on the level of risk in a federal ADM system.
- The Directive addresses systemic bias which does not exist as a factor in existing procedural fairness analysis.

These features of the Directive raise the standard of administrative governance, which is clearly a positive step.[298]

At the same time, the Directive also has gaps and shortcomings. For example:

- The AIA (and resulting category level) appears to focus exclusively on the *impact* of government decision. The AIA's analysis does not appear to include additional factors that courts will consider, such as the expectations of the parties, the specific statutory scheme, and procedure of the administrative body at issue.
- The Directive does not explicitly address an individual's right to a hearing (whether oral or in writing).
- The Directive does not explicitly require reasons, but rather an undefined "meaningful explanation."
- The Directive has significant limitations on its scope and is thus less inclusive than administrative law generally.[299]
- Finally, the Directive does not have the legal status of a statute or regulation and does not create actionable rights for individuals, thus limiting its impact/utility as a tool for vindicating individual rights.[300]

It is possible that some of these gaps and shortcomings will be addressed through updated rules, practices, or litigation. Further, some of these gaps are beyond the jurisdiction of the Treasury Board Secretariat (TBS) and need to be addressed by other government departments. It is important to note that the Directive was created with the requirement it be reviewed and updated for improvement every six months. At the time of this report, TBS was commencing the third review of the Directive. As part of this review, TBS is proposing a number of modifications that expand the scope of the Directive, reinforce transparency and accountability, strengthen protections against discrimination and harm, and clarify requirements. Importantly, the need to clarify "meaningful explanation" is specifically being addressed in this review. [301]

Substantive Fairness

This section will review the current law on substantive fairness and discuss how government AI decision-making challenges these principles.

Standard of Review

When challenging a government decision (made by AI or otherwise), a key legal question is to determine the standard of review i.e. how much deference should the court give to the government decision-maker or regulatory body. As recently clarified in *Vavilov*, there are three potential standards of review:

- For most government decisions, the default standard of review is reasonableness.
- If a matter fits into an exception (constitutional questions,[302] questions of law of central importance to the legal system, questions regarding jurisdictional boundaries) the standard of review is correctness.[303]
- If a governing statue provides for a statutory right of appeal,[304] general appellate standards of review apply,[305] which in most cases is palpable and overriding error.

The LCO believes it is unlikely that government decisions made or aided by AI will be subject to the correctness standard. [306] Lower court interpretations show that legal principles such as privilege count as central to the rule of law,[307] but that interpretation of a specific statutory code does not.[308] Early court decisions also suggest it is unlikely disputes about a specific AI system used by a specific administrative tribunal will be considered "general questions of law of central importance to the legal system as a whole."[309]

In light of this analysis, the LCO believes government decisions made or aided by AI are most likely to be subject to the *Vavilov* reasonableness standard. The Supreme Court explained that "reasonable" is

> *…[t]he principle that the exercise of public power must be justified, intelligible and transparent, not in the abstract, but to the individuals subject to it…*[310]
>
> *…a reviewing court must develop an understanding of the decision maker's reasoning process in order to determine whether the decision as a whole is reasonable.*[311]

The Court further explained that the analysis of the "reasonableness" of a decision is contextual:[312]

> *A reasonable decision is one that is based on an internally coherent and rational chain of analysis and that is justified in relation to the facts and law that constrain the decision maker.*[313]

The factors that constrain the decision maker are:

- The governing statutory scheme
- Other relevant statutory or common law
- The principles of statutory interpretation
- The evidence before the decision-maker
- The submissions of the parties
- Past practices and decisions of the administrative body
- The potential impact of the decision on the individual to whom it applies

Further, in *Vavilov* the Court stated that where reasons are required, they must "provide a transparent and intelligible justification" for the outcome,[314] suggesting that justification is the central part of the test.[315]

The justification requirement applies "both to the process of articulating the reasons and to outcomes."[316] Scholars have interpreted this requirement to mean that a reasonable decision "is a decision that can be explained to follow from a set of facts and premises[317] Put differently, "outcomes must be shown to be justified by clear and compelling arguments."[318]

Applying a *Vavilov* reasonableness analysis without specific facts is difficult. However, the LCO and others have identified AI-related questions that Canadian courts will likely face on judicial review of government AI decisions. These questions fall into three broad categories:

- What are reasons from an AI system?
- Can AI meet the reasonableness standard?
- Who is held to the reasonableness standard?

The answers to these questions are key to ensuring legal accountability for government AI systems.

Before addressing these questions, it is important to discuss AI explainability and the concept of "human in the loop", both of which have important implications for judicial review and legal accountability of government AI systems.

Explainability

"Explainability" is a key concept in AI development and policy-making that is inextricably linked to the legal principles of reasons, intelligibility and justification. Explainability is not a new concept in computer science, but its importance has grown as AI systems proliferate.[319]

Within the technical and legal communities, there is consensus on the significance of, and need for, AI explainability. There is not, however, a commonly agreed-upon definition of explainability or how it can be achieved.

In very general terms, explainability is about making an AI or algorithmic system or decision comprehensible to a human. This concept, while apparently straightforward, is extraordinarily complex. As one commentator notes:

> *There are almost as many forms of explainability as there are AI-related problems that need to be understood and solved. For that reason, explainability can never be separated from the underlying AI-related harm or concern (fundamental rights, usability, safety) that explainability is seeking to address.*[320]

Explainable AI can be categorized several ways, including both general explainability (which explains the functioning of the AI in its entirety) and local explainability (which explains a specific AI decision).

The variables in explainability are endless and there are many challenges in trying to make AI understandable to humans. For example:

- Different audiences may need an AI system to be "explainable" for different reasons. Computer scientists may have much different needs than courts, for example.[321]
- There is considerable variation in attributes that make a system understandable.[322]
- There is variation in what specific terms mean.[323]
- There can be challenges in AI systems translating "explanations" to human-interpretable concepts.[324]

- There are variations in how explainability is presented: text explanation, visual explanation, local explanation, explanation by example, explanation by simplification, or feature relevance explanation.[325]
- There is variation in the goals sought to be achieved by explainability, potentially including trustworthiness, causality, transferability, informativeness, confidence, fairness, accessibility, interactivity, and privacy awareness.[326]

Over time, many of these questions may be resolved. For the foreseeable future, however, "explainable AI" complicates important administrative law issues.

Human-in-the-Loop

Like explainability, human oversight is widely considered to be an important principle to promote fairness and non-discrimination in AI systems. Unfortunately, at present human-in-the-loop issues also complicate important administrative law issues.

Consider the many stages where humans may be involved in a government AI system:

- The person or organization that developed the AI system.
- The government, agency, administrator, or adjudicator who decide how to use the AI system.
- The frontline workers who use the AI system or interpret its recommendations or decisions in individual cases.
- The person or organization that supervises, monitors, updates, and evaluates the AI system.
- The agency, tribunal, or court who reviews the government decision made or aided by the AI system.

The decision as to who, and when, there will be a human-in-the-loop is significant technically, operationally and legally. At this stage, it is not clear what level of human involvement is sufficient to ensure legal accountability/oversight for a government AI system or decision.

AI may significantly reduce the exercise of discretion, even when there is a "human in the loop." AI risk prediction tools may result in "automation bias," or overreliance on machine outputs.[327] A machine-generated decision template might even convert decision drafting to simply signing an automated body of text.[328]

Finally, some commentators have raised questions about a court or tribunal's capacity to assess extraordinarily complex technical and statistical evidence. As a result, some believe that courts and tribunals assessing AI-based decisions are likely to "remain poorly situated to review the accuracy of the machine learning model as a whole."[329]

Reasons from an AI System

Reasonableness review starts with the reasons.[330] Reasons are central to a court's assessment and understanding of how a government decision was made. *Vavilov* raises the standard of quality of reasons to be provided.

What qualifies as reasons under the *Vavilov* standard?

Courts have held that a reasoned explanation must show adequacy, logic, coherence and rationality:

> *The reviewing court must be able to discern an "internally coherent and rational chain of analysis' that the 'reviewing court must be able to trace' and* ***must be able to understand****…"* [emphasis added].[331]

Moreover, the reasons must be intelligible and transparent, "not in the abstract, but to the individual impacted by it."[332]

Explainability models have been suggested as possible "reasons" or "explanations" for an AI system's output.[333] There are at least two difficulties with this suggestion:

First, as discussed above, explainability (the computer science concept) is not the same as justification (the legal concept). For example, an explainability model that assesses whether an AI system is performing as the designers intended may not demonstrate that a decision was coherent and rational.

Second, explainability technical models are varied and complex and whether one qualifies as "reasons" will depend very much on the circumstances and complexity of an AI system. Given that courts have held that reasons need to be justified "not in the abstract, but to the individual impacted by it", it may be that only explainability models that "explain" specific decisions – as opposed to a system generally – would be legally sufficient.

If an AI system's "reasons" are not available or are insufficient, the court could look at the "record" or "context" of the decision.[334] In the context of AI, if a system does not have an explainability model (or if the model is insufficient to meet legal standards), the reviewing court may have to assess how the AI system was designed, its data, and/or testing and validation reports. In some situations, this information, if available and produced, could lead to greater transparency than a decision made without AI. In other situations (decisions made by complex machine learning systems, for example), humans may *never* understand how or why the system produced the result it did.

At this stage of "explainable AI," there are many outstanding technical and legal questions. As a result, tribunals and courts will be faced with many challenging issues to determine what qualifies as "reasons" from an AI system.[335] This is another area where policy guidance would be beneficial to assist tribunals and courts address complex AI issues.

AI and the *Vavilov* "Reasonableness" Standard

The Supreme Court has described reasonableness review as an approach that focuses on justification, offers methodological consistency and reinforces the principle "that reasoned decision-making is the lynchpin of institutional legitimacy."[336] Further, for a decision to be reasonable, a court must be satisfied that "there is [a] line of analysis within the given reasons that could reasonably lead the tribunal from the evidence before it to the conclusion at which it arrived."[337]

Will AI systems meet this standard?

Canadian administrative law places greater significance on reasons and justification as the stakes in a administrative decision increase. In other words, government action that has a significant impact on an individual's life will have a higher expectation for justification. Given the state of the technology and outstanding legal questions, it is possible that some AI systems cannot and will not ever meet this threshold.

This is not to say that the *Vavilov* "reasonable" standard should be adjusted for AI-based government decision-making. On the contrary, governments and administrative bodies should consider whether some regulatory actions or government decisions are not suitable to be made or assisted by AI systems. This is especially true where an important government decision is made entirely by an AI system without a "human-in-the-loop." If an AI system cannot be understood by a court, the court will have to decide whether the opacity of the AI system renders the decision inherently "unreasonable."

Concluding Thoughts on AI and Administrative Law

Administrative law will have a profound impact on government use of AI, ADM and related technologies. These systems will have to be designed, administered and evaluated to ensure compliance with two dimensions of administrative law: procedural fairness and substantive fairness. Unfortunately, as noted by administrative law expert Jennifer Raso, "administrative law is more gap than law when it comes to algorithmically-driven decision-making."[338]

For example, procedural fairness requirements apply to AI systems when they impact an individual's rights, interests, or obligations. Yet, there are many examples of AI systems deployed in government decision making that undermine or ignore procedural fairness obligations.[339] The examples discussed in this paper show that challenging an AI system in court for lack of procedural fairness can result in a system being re-designed[340] or abandoned altogether.[341]

The sheer number of regulatory tribunals in Ontario, the variety of functions they serve, and the myriad individual rights and interests that can be affected make procedural fairness a complex area of law. The principles of procedural fairness (fairness, transparency, accountability) and the safeguards created to achieve these principles (notice, participation, impartiality, reasons, rights of appeal) are the same for all administrative decisions, but how they are applied differs depending on the regulatory body, their governing legislation, and the decision in question.

Similarly, in assessing the substantive fairness of a regulatory decision, courts will have to consider what qualifies as reasons from an AI system. Can the output of an AI system show an internally coherent and rational chain of analysis?[342] Will machine generated "explainability" meet the legal standard of "reasonableness"? Is it possible to know whether the reasons generated by an AI system describe the actual justification for the decision?

The Canada ADM Directive is a good start to addressing these issues but is not a complete solution, even at the federal level. Many features of the Directive raise the standard of administrative governance, which is clearly a positive step. However, the Directive also has gaps and shortcomings, even within the realm of federal administrative law.

Importantly, the Directive provides high level guidance for administrative agencies to adapt to their own circumstances. The extent to which procedural and substantive fairness is achieved at the federal level will be determined by the details of how each system is designed, built and executed. It is possible that some of the gaps and shortcomings of the Directive will be addressed through updated rules, practices or litigation.

The challenge of reconciling AI, ADM and related systems with administrative law requirements will be much greater in Ontario. This is because there is no equivalent of the Canada ADM Directive at the provincial level, nor are there equivalent policies or directives at the municipal or provincial agency level.

Absent a Directive-like instrument at the provincial level, courts and tribunals will be compelled to

answer a complex series of questions on a case-by-case basis, without the benefit of comprehensive technical or legal guidance, including but not limited to:

- What constitutes "notice" when an AI system is used by the provincial government, a municipality or provincial agency?
- Do impacted parties need to know only that an AI system was used in the process of the decision, or do they need more information about how the system works, such as what data it relied on, how the data was sorted and weighted, or whether a human was involved in the decision at a certain stage?
- How can the provincial government, municipalities or agencies assess the "reasonableness" of a decision made or influenced by AI?
- What are "reasons" from an AI system?
- How does "explainable AI" accord with the legal "justification" for a regulatory decision?
- How should courts and tribunals assess risk and impact of a government AI decisions?
- Can an AI system be explained sufficiently so that affected parties understand the decision made against them?
- Are meaningful appeal options possible?
- How do parties participate in a decision made in part or in whole by an AI system?
- Does procedural fairness require that parties be consulted in the creation of a system?
- How do parties navigate these disputes in an efficient and cost-effective way?

Coherent policy guidance on these issues is needed at the provincial level. Otherwise, the use of AI by governments to assist decision-making brings great risks for the individuals affected by these systems, for the courts and tribunals reviewing these decisions, and for the governments and agencies developing and operating AI systems.

For the individuals affected by these systems, the lack of clear policy guidance risks incorrect or illegal denials of government benefits, licenses, business opportunities, etc. This is because, absent appropriate protections, the access to justice and legal protections/remedies that should be available to all Ontarians will likely be only available to the best resourced litigants.

For courts and tribunals reviewing government AI decisions, the lack of policy guidance means that triers of fact and decision-makers will be compelled to assess every complex legal and technical issue on a case-by-case basis. Addressing these issues in this manner will result in delays, additional court resources, and poor decision-making. Case-by-case decision-making also means courts cannot give clear legal guidance to developers, government officials, tribunals/courts, legal professionals and litigants.

Finally, for governments and agencies developing and introducing these systems, the lack of policy guidance creates a risk that systems will deliver poor services, are biased, are subject to litigation challenges, and lack public credibility.

The LCO believes it is clearly preferable for policymakers in Ontario to proactively address these issues through some form of legislative or policy guidance. The framework discussed in the LCO's *Regulating AI* Issue Paper builds on and advances the Canada ADM Directive. This model would go a long way to meeting the procedural and substantive fairness required by Canadian administrative law. Equally important, this framework would mitigate the cost, confusion, and time of litigating these issues on a

case-by-case basis over many years. Finally, this framework would create a transparent, consistent, and level and playing field for public service administrators, technologists, tribunals, courts, litigants and the justice system.

Accordingly, the LCO recommends that the Government of Ontario develop and adopt an Ontario AI Directive to guide provincial AI development and serve as template for other public organizations under provincial jurisdiction.

A provincial directive need not address every potential administrative decision. One of the purposes of this instrument could be to delineate when administrative law obligations *are not* required due to the nature of the decision at issue.

A provincial directive could be followed up by guidance tools that address more specific issues. This approach would have many benefits, including:

- Promoting procedural fairness in the design and implementation of government AI systems.
- Providing designers and operators of AI systems with greater clarity as to their legal obligations.
- Promoting legal and administrative consistency within and between government departments and public agencies.
- Providing important direction to courts considering procedural reviews of government decisions.
- Decreased costs and delays in the justice system.
- Promoting legal accountability and access to justice.

Readers should note there is a growing body of academic thought about the long-term implications of AI on administrative law and administrative decision-making in the U.S.[343] Danielle Citron, one of the pioneers in this area, is skeptical about whether judicial review of complex AI systems is realistically possible. As a result, she and others have suggested the need a new model of "technological due process."[344] Some commentators are more optimistic and suggest that AI systems will improve legal accountability and government decision-making, perhaps dramatically. Still others caution against what they call "techno-utopianism."[345] In this view, administrative law will have to be adapted significantly meet the new realities of AI-powered government decision-making.[346]

The LCO agrees there are reasons to be cautious. Government algorithmic and AI systems to date have a mixed track record. That said, there also are many reasons to be optimistic about AI in government decision-making and to encourage its use. This optimistic future will not create itself, however, and is dependent on the choices made in the relatively early stages of government AI systems.

In order to ensure government AI systems comply with administrative law requirements, the LCO makes the following recommendations:

Recommendation 13

The provincial government should develop and adopt an Ontario AI and ADM Directive to guide provincial decision-making and serve as template for other public organizations under provincial jurisdiction. A provincial directive need not address every potential administrative decision. A provincial directive could be supplemented with guidance tools that address more specific issues. Factors that the Directive could address include:

- **What constitutes "notice" when an AI system is used by the provincial government, a municipality or provincial agency?**
- **Do impacted parties need to know only that an AI system was used in the process of the decision, or do they need more information about how the system works, such as what data it relied on, how the data was sorted and weighted, or whether a human was involved in the decision at a certain stage?**
- **Do parties need to participate in a decision made in part or in whole by an AI system?**
- **Do parties need to be consulted in the creation of a system?**
- **What type of "reasons", "jusitification" or "explainability" is required?**
- **How to assess the "reasonableness" of a decision made or influenced by AI? What is the standard an AI system will be held to?**
- **Can an AI system be explained sufficiently in a way that impacted parties can understand the decision made against them and are meaningful appeal options possible?**

Recommendation 14

The provincial government, community legal clinics, members of the private bar, academics, the judiciary, tribunals, and technologists should come together to consider the following issues:

- **How to determine the evidence required when assessing a government AI system?**
- **How to assess the "reasons" of a government AI system?**
- **How to assess the risk and impact of a decision made or aided by a government AI system?**
- **How to ensure a there is a "human-in-the-loop" while protecting against automation bias?**
- **Is a government AI system sufficiently understandable to meet the requirements of justification?**
- **Does AI "explainability" models provide transparent and intelligible understanding of how a specific outcome was reached?**
- **How can appeals from decisions of AI systems meaningful to the parties and the**
- **Who is to be held to the standard of reasonableness?**
- **How can parties, courts and tribunals address challenges to government AI systems in a fair, efficient and cost-effective manner?**

Beyond the immediate questions addressed so far, many academics have also begun asking questions about the long-term future and impact of AI on administrative law and legal accountability for government decision-making. These scholars are questioning whether AI requires a fundamental shift in administrative law principles or much higher standards of governance.

In the not too distant future, hard questions will need to be asked about whether Canadian administrative law principles remain viable in a more technologically advanced administrative state. Will the comparatively comprehensive Canada ADM Directive still be a viable tool when the speed, scale, and number of AI systems in government expands rapidly, as it is likely to do? In those circumstances, do the procedural fairness protections in the Directive become more important or less? Will compliance with the Directive become easier or more difficult?

One of the best summaries of the long-term impact of AI on administrative agencies, and administrative law, is a 2020 report titled *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*.[347] In a section titled "Implications: The Future of Mass Adjudication" the authors lay out the opportunities, and risks, of AI and government administration succinctly:

> *At its most ambitious, AI could transform what it means to adjudicate a case. To be sure, current use cases are a far cry from full automation of adjudication, but the trajectory raises profound implications for the [future of] adjudication…*
>
> *First, the trajectory of AI tools in adjudication raises the normative question about the desired extent of discretion in adjudication. [The Social Security Administration] moved early to formalize its policy as a decision tree…More rules-based adjudication may promote consistency, but may also undercut one of the rationales for adjudication: tailoring the application to individualized circumstances...*
>
> *Second, the development of AI tools raises questions about notice and transparency. Formal adjudication requires that a decision be based on the exclusive record, but AI tools involve a transfer of decision-making authority away from line-level adjudicators toward AI developers…At the state level, where benefits programs have effectively modified eligibility criteria through the use of algorithmic decision-making, some courts have found that the change violates notice and comment requirements and deprives claimants of due process…*
>
> *Third, the adoption of AI tools could potentially erode the decisional independence of and de novo review by ALJs [Administrative Law Judges]. An SSA ALJ has a "duty to fully and fairly develop the record and to assure that the claimant's interests are considered." The duty is heightened where the claimant is not represented by counsel but exists also with represented claimants…[There has been] pushback by ALJs to perceived infringement on their decisional authority, but such pushback may weaken if adoption is seen to ease the work burden. Automation bias could mean that ALJ review of AI-generated content becomes increasingly perfunctory… This dynamic of overreliance may be particularly acute given the high caseloads that adjudicators face*
>
> *Last, as we have seen in many areas of machine learning, the adoption of such tools can heighten concerns of bias. …*
>
> *Forecasting the trajectory of AI tools brings into relief longstanding debates about the core values of agency adjudication. At its best, AI may address longstanding problems of the accuracy and consistency of decisions. By increasingly automating core portions of the adjudicatory process, these tools may cut down on staggering agency caseloads without a sacrifice in the accuracy of decision- making. At the same time, this future of algorithmic adjudication may cause us to go back to the basic premises of procedural due process. Why do we hold hearings? Machine learning may enable agencies..to expedite decisions by skipping resource-intensive hearings.*

> *And while this may meet the goals of accuracy under due process, it may also cause us to revisit the lost constitutional rationale of dignity. The rationale for hearings may not solely be to promote accuracy, but also to explain the law, to engage claimants, and to make them feel heard. This, then, is the challenge of the push for AI solutions in mass adjudication: Agencies seek out these solutions to accelerate case processing, but that same pressure may cause agencies to crowd out the dignitary values of an adjudicative hearing.*[348]

Danielle Citron and Ryan Calo suggest the question is not "how to restore the status quo ex ante given that machines have supplanted people…[but rather] whether technology obligates a fundamental re-examination" of why administrative bodies are bestowed with decision-making power in the first place.[349] In their view, "[a]gencies that automate throw away expertise and discretion with both hands."[350]

From a Canadian perspective, Teresa Scassa points out that the factors justifying the creation of administrative bodies are quite different than the factors justifying the automation of government decision-making. In her view, administrative bodies were developed because of:

- The desire to depoliticize certain decisions.
- The need for greater specialization and technical or subject matter expertise.
- A reluctance to enmesh courts in matters not suitable to judicial review.
- Because many individuals or organizations subject to government decision-making (such as trade unions) viewed the courts as antithetical or unhelpful.[351]

By way of contrast, the most common justifications for automating government decisions are to reduce costs, enhance speed and efficiency, and improving the quality of fairness.[352] If automation and AI shift the purpose of administrative bodies and tribunals, do we need to fundamentally reconsider the principles that govern administrative decisions?

At this stage of government AI development, these questions are impossible to answer. Nevertheless, it is difficult to conceive of a future where the principles of transparency, participation, impartiality, fairness, and reasons are not central to ensuring the legal accountability of government decision-making. It is also true that the significance of these principles will always be heavily dependent on context and impact. Administrative will have to adapt accordingly over time.

Privacy and Data Protection

Privacy is a seminal issue in discussions about AI and other data driven technologies.[353]

The tension between protection of individual privacy rights and the benefits of AI innovation was quickly recognized as a significant issue in the use and development of AI technologies. This tension is particularly potent in the public sector where governments hold copious amounts of personal data including, but not limited to, health records, property ownership and tax records, employment information, birth, school, marriage and death records, social benefit information and criminal records.

Government policy decisions and services are improved when governments can access and process data. For example, ICES, the not-for-profit research institute that uses analytics to improve health care in Ontario, [354] was created twenty-five years ago as an exception to privacy legislation, specifically in recognition of the valuable work that can be done when researchers have access to useful data.[355]

Privacy and data protection are often discussed interchangeably, but they are distinct concepts. Privacy is a broad legal concept that is considered significant for personal physical and moral autonomy and well-being.[356] There is "a common understanding that control over one's own personal information is central to a self-determining and responsible being."[357] Privacy legislation in Ontario and Canada protects a person's right to their personal information. Data protection, on the other hand, addresses the safety and security of data. Privacy and data protection both include rules relating to the collection, use and disclosure of personal information.

The LCO begins this discussion with several introductory points points:

- Data protection legislation is divided in Ontario between public, health and private sectors. The focus in this paper is public sector data protection.
- Many governments are seeking to amend public sector laws to allow for more intra-governmental data sharing.
- The combination of AI and greater intra-governmental data sharing raises important privacy, accuracy, bias and surveillance issues. For example, policymakers, stakeholders and (eventually) courts will need to consider how to balance government efforts to reduce fraud in government benefits with the need to protect privacy. [360] While AI may assist in the detection of fraud, it could also raise concerns about fishing expeditions or investigations against those who are falsely identified.[361] This raises important questions about AI as a surveillance technology.

These developments raise questions about whether there are sufficient protections in Ontario's legislative framework to ensure Ontarians understand who is accessing their data and how it is being used.

Ontarians can learn from the examples in Australia (Robodebt), Michigan (MiDas), and the Netherlands (SyRi). All three systems involved AI that scanned huge amounts of government data to find "irregularities" in either individual employment insurance records (Robodebt/MiDas) or social benefits determination (SyRi). All three systems were widely criticized, legally challenged, and eventually reformed or cancelled due to privacy, due process or system accuracy concerns. All three systems also caused great harm (erroneous denial of benefits, freezing assets, conducting further investigations without notice or explanation, loss of employment, loss of homes and loss of life). The repercussions for each government was significant. In Australia, for example, the government lost a US $1.7B class action lawsuit. In Michigan, the government lost a similar class action suit. And the Dutch government was ordered to shut down the welfare fraud system, and the childcare benefit system controversy caused the government to resign.[362]

Ontario Privacy Amendments

In the last few years, the provincial government has amended laws that facilitate data sharing and usage. These changes support the growth and development of government AI systems. Many of these changes raise important questions about how privacy will be protected in provincial AI systems in the future.

Ontario's *Freedom of Information and Protection of Privacy Act (FIPPA)*[363] protects personal information held by government organizations and provides individuals with a right of access to their own personal information. FIPPA applies to "the collection, retention, use, disclosure and disposal of personal information" in the government's "custody or control."[364] *FIPPA* was enacted over thirty years ago and has been amended frequently. Most recently, in 2019 the Government of Ontario enacted two significant changes to address AI data governance:

The first amendment allowed for "data integration" between government departments.[365] Sharing of data among government departments challenges individual privacy protections because individuals may not know who has access to their information, what purpose it is being used for, or how they may be affected if their information is shared. The amendments address these concerns, in part, by introducing compliance obligations that include "the establishment of processing standards, de-identification to the extent possible, and oversight and approval of procedures by the IPC."[366]

The second amendment broadens the government's ability to disclose data for investigatory purposes.[367] The language expands the government's data disclosure authority from "disclosure to aid with an investigation where a law enforcement proceeding is likely to result" to include disclosure permitted where there is a "reasonable basis to believe that an offence may have been committed".[368] Privacy experts are concerned this change "expand[s] the ability of public sector bodies to share personal information with law enforcement without consent."[369]

In March 2020, the province enacted further amendments as part of an omnibus bill largely seen as the first provincial response to the COVID pandemic.[370] These changes permit sharing of government data with extra-ministerial parties, including both profit and not-for-profit organizations.[371] Privacy experts cautioned that the changes were unnecessarily broad and could result in "the creation of significant databases, derived from public sector data, held and controlled by private sector entities."[372]

Provincial Data Strategy

Coinciding with *FIPPA* amendments were several provincial reforms to data governance. These reforms demonstrate the province's commitment to expanding government AI in Ontario. Like the *FIPPA* amendments described above, however, these reforms raise questions about privacy protection in provincial AI systems.

The Province began these initiatives in 2019 with the release of the *Ontario Data Strategy* initiative.[373]

In April 2021, the province released its *Digital and Data Strategic Directive* which was described as "bringing major change to government by embracing technology and harnessing innovation."[374] The Directive included an announcement to develop the province's Artificial Intelligence Framework.

The *Directive* includes generalized assurances about privacy protections in "ministry technology projects."" More specifically, the Directive states that it will

> *… ensure information and data-related risks, design considerations and requirements such as privacy protection, information security, access and recordkeeping are integrated into the design of ministry technology projects and procurements and are subsequently operationalized and monitored.*[375]

The province is also creating a data authority that will be tasked with building the modern data infrastructure necessary "to support economic and social growth," while ensuring that this information is private, anonymized, and securely stored and handled.[376]

Finally, the province introduced the Ontario Health Data Platform (OHDP) which is designed to be a high-performance, accurate, privacy protection mechanism that links large health data sets held by multiple organizations across Ontario "to allow for big data analytics, including machine learning, that will strengthen evidence for Ontario's ongoing response to COVID-19 and its related impacts."[377]

These developments support the provincial government's commitment to AI growth. It remains to be seen, however, how individual privacy rights will be protected when personal information is being shared and how personal information will be used in the future.

Broad Implications of Provincial Privacy and Data Governance Reforms

These developments and other initiatives in Ontario, Quebec and at the federal level have focused public attention on AI, privacy law, and data protection.

Courts, policymakers, and the public will have to decide how to effectively balance individual privacy rights with the potential benefits of government AI systems that rely on personal data.

As the province moves forward, there are gaps that should be monitored and addressed:

- The effectiveness of de-identification may be limited. Anonymization of data is sometimes considered an antidote to concerns about privacy. However, AI systems can often re-identify individuals.[378]
- AI challenges an individual's right to access their personal data. For example, are individuals entitled to data that is shared intra- or extra-ministerially or data that is de-identified and re-identified? Further, how will the provincial government ensure personal data information is provided in a format that is understandable?

- What remedies are available for individuals whose personal information has been wrongly shared?
- How can privacy and data governance policymaking be transparent and participatory?[379]

There is a need for public discussion about the balance between the potential benefits of greater government data sharing (improved government services, for example) and the potential risks of data sharing (surveillance, privacy violations, human rights, legal fairness).

This report is not about privacy protections, so the LCO is not making formal recommendations in this area. That said, the LCO is sympathetic to the recent comments of the provincial Information and Privacy Commissioner who stated that a complete review of public sector privacy law needs to be undertaken by the government:

> *The IPC has repeatedly emphasized the need to update our public sector privacy laws to address [technological changes]. They continue to fall behind rapidly evolving digital technology and data-rich information practices. Technology available today has many benefits for society … However, the use of data and technology must not come at the expense of privacy…. While Ontario's legislation has had some updates, including the recent data integration amendments, a comprehensive review is needed to ensure privacy rights of Ontarians continue to be protected in our changing environment. Other provinces have strengthened their laws to meet new privacy challenges.*[380]

Civil Procedure and Evidence

In Ontario, civil cases are governed by procedural rules to ensure proceedings are fair for both sides.[381] So far, this report has discussed gaps and concerns that exist when AI intersects with substantive areas of law like human rights and administrative law. Less discussed, but equally important, are the procedural rules and laws that ensure cases can be presented to a trier-of-fact with fair and appropriate procedure.

The issues discussed in this section include an analysis of the *Rules of Civil Procedure*, confidentiality and sealing orders, Crown privilege and liability, and evidence. At present, the LCO does not believe law reform measures are needed in these areas. Nevertheless, AI-based civil proceedings are likely to be lengthy and costly, raising concerning questions about access to justice for all but the best-funded litigants. Over time, reforms may be necessary to address these challenges.

Rules of Civil Procedure

The *Rules of Civil Procedure* apply to any case brought in the Superior Court and higher. A party challenging the use of an AI system or algorithm for an alleged breach of *Charter*, privacy or statutory breach are likely to bring the matter to the Superior Court in Ontario.[382] Not surprisingly, these proceedings are likely to be lengthy and costly, raising concerning questions about access to justice for all but the best-funded litigants.

Under the *Rules of Civil Procedure*, parties are required to produce "every document relevant to any matter in issue…"[383] In many cases, an AI system's source code, data, variables or other information will be sought by the challenging party, potentially resulting in massive disclosure and overwhelming documentation production. In response, defendants (including governments) would likely argue overbreadth, vagueness or proportionality. In these circumstances, courts will need to balance the potential for a "fishing expedition" against the legitimate evidential needs of the party and trier of fact. In some cases, the creators, owners or users of AI systems will object to productions based on national security, confidentiality or government discretion.[384] Disclosure proceedings are likely to be complex and costly.

Assuming disclosure of this evidence is provided or ordered, counsel and parties will then likely find themselves with perhaps millions of lines of data, source code or other materials. In these circumstances, the disclosure could be both voluminous and understandable only by computer scientists. These issues will be worsened, or perhaps insurmountable, if there is a legal challenge to a complex machine learning or neural network AI system, where system's decision-making process may not be traced or understood by humans.

Then there is the issue of who to examine. A party is entitled to examine one representative of a corporation,[385] or a single Crown witness is designated by the Deputy Attorney General.[386]

A plaintiff challenging an AI system will need to determine who to examine: a representative of the company or group that designed the system (unlikely to be a single person); a person who obtained and manages the system's data; the person who administers the system; the person who ran validation testing and audits; or a senior executive who may have some surface knowledge about all of them – if these people are still available to be discovered at all.

It may be that the *Rules of Civil Procedure* are sufficiently adaptable to allow parties to navigate a AI dispute fairly. However, the novelty and complexity of AI systems suggests the need for extensive education for both counsel and judges.

Early cases in Ontario and Canada will be significant for setting precedent. However, these precedents may be easy to distinguish, given the variety of AI systems and heavily contextual nature of government decisions.

Over time, it may be advisable to amend the *Rules* to clarify or simplify AI litigation (establishing rules about productions or parties to examine, for example). These reforms could depend on how the caselaw develops, where issues arise, and whether they can be corrected or clarified with a *Rule*.

Confidentiality and Sealing Orders

Owners or users of AI systems may object to disclosure of source code or data they have spent time and resources developing. In the U.S., the defense of trade secrets is frequently raised in response to parties seeking disclosure about algorithms and AI systems.[387]

The law of trade secrets is different in Canada, as there is no equivalent to the US *Trade Secrets Act*. Trade secret law in Canada is based on common-law principles and arises through common law torts, such as of breach of confidence or breach of fiduciary duties.[388] In Canada, a party wishing to keep information about an AI system confidential would likely seek a confidentiality or sealing order. The LCO is not aware of a precedent in Canada where a party has sought or contested a sealing order relating to an AI system.

The authority for a sealing order comes from the Ontario *Courts of Justice Act*[389] and is an exception to the open court principle.[390] The Court of Appeal for Ontario recently stated: "sealing orders and other kinds of orders that restrict public access to the court and court proceedings are very much the exception."[391] A party wishing to limit public access to corporate materials is not likely to produce documents on discovery without first obtaining a sealing order.

The test for whether public access to a court file should be restricted is set out in the Supreme Court of Canada's decision in *Sierra Club of Canada v. Canada (Minister of Finance)*.[392] A sealing order should only be granted when:

> *(a) such an order is necessary in order to prevent a serious risk to an important interest, including a commercial interest in the context of litigation because reasonably alternative measures will not prevent the risk; and*
>
> *(b) the salutary effects of the confidentiality order, including the effects on the rights of civil litigants to a fair trial, outweigh its deleterious effects, including the effects on the right to free expression, which in this context includes the public interest in open and accessible court proceedings.*[393]

The LCO does not believe there is anything particular about AI systems would require a new approach to the law of sealing orders. Nevertheless, motions for confidentiality orders could lengthen cases and increase expenses, creating further access to justice challenges.

Crown Privilege and Crown Liability

Parties wanting to challenge a Government of Ontario AI system may need to respond to defences based on Crown privilege or limitations to Crown liability.

Many people within governments are concerned that too much AI system transparency raises the potential that malevolent actors will try to "game the system".[394] The validity of this concern is debatable and has been considered in-depth by legal scholars at NYU and McGill who have concluded that for many systems gaming the algorithm is not feasible and not a true concern.[395] Nevertheless, in certain circumstances, the Crown may raise a "public interest" defense to deny disclosure of important elements of an AI system.[396]

The recently enacted provincial *Crown Liability and Proceedings Act*, 2019, S.O. 2019 (*CLPA*) raises important questions about whether an individual might be barred from bringing a claim of negligence against the Government of Ontario for developing, implementing, deploying and relying on an AI system. The *CLPA* addresses Crown liability in negligence actions.[397] Although issues of AI and tort liability are beyond the scope of this paper, the *CLPA* is mentioned here because it raises access to justice concerns.

Section 11(5) of the new *Act* appears to broaden the areas of provincial government's immunity by expanding the definition of "policy decision"[398] to describe what have historically been understood to be operational decisions and actions.[399] The wording of these sections suggest it is possible that parties will not be able to sue the provincial government for negligent government activity, which could potentially include systemic institutional wrongs.

To date, courts in Ontario have shown a reluctance to interpret the legislation broadly.[400] In *Francis v. Ontario,*[401] the Ontario Court of Appeal said that adopting an expansive meaning of section 11(5) would "directly offend the purpose behind the statutes limiting Crown immunity." Further, "[e]xempting all government actions from liability would result in intolerable outcomes."[402]

It is likely that a government decision to develop an AI system would be a policy decision and not subject to judicial scrutiny. A decision to outsource the creation and design of that system could also be a policy decision. The LCO suggests, however, that the deployment, use and maintenance of an AI system should likely be considered "operational" and thus not barred from negligence claims. Expanding the use of AI in government decision-making introduces new potential harms to individuals on a wide scale. It will be important to maintain the potential of tortious liability when experimenting with these new technologies.

Evidence

Evidentiary hurdles and complications could arise in cases where parties are disputing the use of an AI system. For a plaintiff who needs to prove on a balance of probabilities that the use of an AI system is in breach of a law, or a defendant trying to justify its use, extensive information about the system will likely be required.

Even if all available information about the AI system is disclosed, some AI systems will still lack interpretability. As described by Professor Ashley Deeks:

> *...because a machine learning system learns on its own and adjusts its parameters in ways its programmers do not specifically dictate, it often remains unclear precisely how the system reaches its predictions or recommendations. This is particularly true for "deep learning" systems that use "neural networks," which are intended to replicate neural processes in the human brain."*[403]

As a result, the "black box" nature of AI systems and algorithms could create a highly burdensome and potentially impossible evidentiary threshold for a plaintiff to meet.

When litigating an AI system – particularly anything more complex than a very basic algorithm - the proceedings risk becoming a battle of experts. Depending on the issue at hand, experts could be required to explain the source code of an AI system, the data the system relies on, the training method, the factors and their weighting, scoring, validation, efforts at correction for bias and discrimination, maintenance, and explainability of a particular system.

Courts have long experience grappling with the challenges and complexities of expert witnesses. Concerns with expert evidence are well articulated.[404] Courts have developed methods to address these concerns, including the *Mohan* test which lays out four criteria for admissibility of expert evidence: (i) relevance, (ii) necessity, (iii) the absence of any exclusionary rule, and (iv) that the witness was a properly qualified witness. Courts will also undertake a cost benefit analysis and experts are required to be independent, impartial and unbiased.[405]

Admissibility and weight given to expert evidence is an important aspect of the justice system[406] and an evolving area of law.[407] AI litigation will be context specific and there are potentially many ways for triers-of-fact to evaluate an AI system. Moreover, the variation in AI systems and the fact they are constantly evolving could lead to novel science being introduced in each new proceeding, limiting their precedential value. In these circumstances, a key question will be whether the current law evaluating expert evidence is sufficient and effective when applied to AI litigation or whether gaps emerge that need to be addressed. This is an important question that should be examined as the use of government AI systems increase.

A related issue in legal practice is the introduction of AI as evidence in the courtroom. There are many examples of this in the criminal justice system.[408] The use of AI as evidence raises many questions including whether it should be judged by the standard of direct witness testimony, expert witness testimony, or measurement using established technology.[409] Areas of particular challenge with introducing AI as evidence in a court room are relevance and authenticity.

Taken together, the potential evidential challenges posed by government AI systems are significant. As described by Gerald Chan and Mable Lai:

> *These novel technologies will demand a significant degree of upfront work and planning in order to use the resulting evidence to its full and proper potential, to challenge evidence with the appropriate vigor, and to flush out unreliable evidence where appropriate. The effective and even-handed presentation of algorithmically generated evidence presents one of the most exciting substantive and practical challenges for criminal and civil litigations in the modern technological era.*[410]

As noted above, the LCO does not believe specific law reform measures are needed in the areas of civil procedure, confidentiality orders, Crown liability or the laws of evidence at the present time. In the short term, the issues identified in this section will be best served by implementing Recommendations 1-14 in this report. In the long-term, further reforms may be necessary to address the challenges addressed in this section.

In addition to Recommendations 1-14 above, the LCO recommends:

Recommendation 15

Ontario's Rules of Civil Procedure should be monitored. As the law develops, new AI-specific Rules of Civil Procedure should be considered.

Recommendation 16

The provincial government should not be immune to tortious liability for government AI systems. The new provincial Crown Liability and Proceedings Act, 2019 should not be used to bar negligence claims against the provincial government for developing, implementing, deploying and relying on AI systems.

Recommendation 17

Ontario's laws of evidence should be monitored to gauge whether the current law evaluating expert evidence is sufficient and effective when applied to AI litigation.

Recommendation 18

The provincial government, judiciary, court administrators and provincial legal organizations should develop educational programs and materials for the judiciary, tribunal members, counsel and administrators.

Recommendation 19

The development and use of AI in Ontario's justice system should be monitored. The provincial government, judiciary, academics, NGOs, and legal organizations should consider establishing a working group or measures to analyze, monitor and report on the use of AI and algorithms in Ontario's civil and administrative justice systems.

Appendix A

Recommendations

Trustworthy AI

To support Trustworthy AI in Ontario, the LCO recommends:

1. The provincial government should not deploy high-risk AI or automated decision-making technologies prior to adoption of its comprehensive Trustworthy AI Framework.
2. The Trustworthy AI Framework should be established in legislation and regulations.
3. The Trustworthy AI Framework should promote AI transparency, accountability, and public engagement in the development, operation, and evaluation of provincial AI systems.
4. The provincial government should create an AI framework to specifically address AI systems that are developed, or used in, the criminal justice system, such as facial recognition, biometric identification, predictive policing and bail/sentencing risk assessments.
5. The Trustworthy AI Framework should establish a framework for municipalities, provincial agencies, and courts and tribunals under provincial jurisdiction.
6. The provincial government should commit to assisting municipalities and public agencies develop resources, tools, and standards to ensure Trustworthy AI in these organizations.
7. The provincial government should develop public performance metrics to ensure the province is meeting the goals of Trustworthy AI.
8. The provincial government should establish a multidisciplinary Trustworthy AI Expert Advisory Task Force and public consultation plan to advise provincial policymakers on how to fulfill the commitments and recommendations herein.
9. The provincial government should continue to seek meaningful and multidisciplinary public input and participation in all phases of AI regulation development.

AI Regulation

To ensure government AI is properly regulated, the LCO recommends:

10. The provincial government's Trustworthy AI Framework should be established in legislation and regulations. The legislation should include, but not be limited to, provisions to ensure provincial AI, ADM and related systems are transparent, accountable, and legal. Legislation should also include provisions that promote access to justice, address bias/discrimination, and a requirement to mitigate harms.

The comprehensive regulatory regime should include:

11. Baseline requirements for all public sector AI, ADM and related systems, irrespective of risk.
    - Strong protections for AI and ADM transparency, including disclosure of both the existence of a system and a broad range of data, tools and processes used by the system.
    - Mandatory "AI Registers".
    - Mandatory, detailed and transparent AI or algorithmic impact assessments, including the identification of prohibited and high-risk systems.
    - Explicit compliance with the Charter, human rights legislation and administrative law.
    - Explicit requirements to measure, correct and audit/monitor bias in AI systems.
    - Data standards.
    - Access to meaningful remedies.
    - Mandatory auditing and evaluation requirements.
    - Independent oversight of both individual systems and government use of AI, ADM and related systems generally.

Human Rights and AI

To ensure government AI systems comply with human rights requirements, the LCO recommends:

12. The provincial government, Ontario Human Rights Commission, technologists, human rights experts, and community members work together to develop a provincial human rights strategy for Government of Ontario AI systems. This strategy should address the following issues:
    - A made-in-Ontario AI Human Rights Impact Assessment
    - Data standards
    - Evidential standards in government AI systemic discrimination cases
    - Guidelines or metrics to measure bias and discrimination in government AI systems
    - Bias testing or auditing requirements
    - Determining reasonable accommodation in government AI systems
    - Reviewing remedy provisions in the Ontario Human Rights Code for sufficiency to address potential harms of government AI systems
    - Access to justice challenges

    A key element of this strategy should be to develop guidance for policymakers to determine what AI systems or applications should be prohibited on human rights grounds.

    This strategy should require:

    - That human rights experts and communities to be involved in the design, development and operationalization of government AI systems
    - That human rights experts and communities to be meaningfully engaged throughout the lifecycle of a government AI system

Administrative Law

To ensure government AI systems comply with administrative law requirements, the LCO recommends:

13. The provincial government should develop and adopt an Ontario AI and ADM Directive to guide provincial decision-making and serve as template for other public organizations under provincial jurisdiction. Factors that the Directive could address include:
    - What constitutes "notice" when an AI system is used by the provincial government, a municipality or provincial agency?
    - Do impacted parties need to know only that an AI system was used in the process of the decision, or do they need more information about how the system works, such as what data it relied on, how the data was sorted and weighted, or whether a human was involved in the decision at a certain stage?
    - Do parties need to participate in a decision made in part or in whole by an AI system?
    - Do parties need to be consulted in the creation of a system?
    - What type of "reasons" or "explainability" is required from a system?
    - How to assess the "reasonableness" of a decision made or influenced by AI? What is the standard an AI system will be held to?
    - Can an AI system be explained sufficiently in a way that impacted parties can understand the decision made against them and are meaningful appeal options possible?

14. The provincial government, community legal clinics, members of the private bar, academics, the judiciary, tribunals, and technologists should come together to consider the following issues:
    - How to determine the evidence required when assessing a government AI system?
    - How to assess the "reasons" of a government AI system?
    - How to assess the risk and impact of a decision made or aided by a government AI system?
    - How to ensure a there is a "human-in-the-loop" while protecting against automation bias?
    - Is a government AI system sufficiently understandable to meet the requirements of justification?
    - Does AI "explainability" models provide transparent and intelligible understanding of how a specific outcome was reached?
    - How can appeals from decisions of AI systems meaningful to the parties?
    - Who is to be held to the standard of reasonableness?
    - How can parties, courts and tribunals address challenges to government AI systems in a fair, efficient and cost-effective manner?

Civil Procedure, Evidence and Other Issues

In addition to Recommendations 1-14 above, the LCO recommends:

15. Ontario's *Rules of Civil Procedure* should be monitored. As the law develops, new AI-specific *Rules of Civil Procedure* should be considered.
16. The provincial government should not be immune to tortious liability for government AI systems. The new provincial *Crown Liability and Proceedings Act*, 2019 should not be used to bar negligence claims against the provincial government for developing, implementing, deploying and relying on AI systems.
17. Ontario's laws of evidence should be monitored to gauge whether the current law evaluating expert evidence is sufficient and effective when applied to AI litigation.
18. The provincial government, judiciary, court administrators and provincial legal organizations should develop educational programs and materials for the judiciary, tribunal members, counsel and administrators.
19. The development and use of AI in Ontario's justice system should be monitored. The provincial government, judiciary, academics, NGOs, and legal organizations should consider establishing a working group or measures to analyze, monitor and report on the use of AI and algorithms in Ontario's civil and administrative justice systems.

ENDNOTES

1 Danielle Keats Citron, Technological Due Process (2008) 85 Wash U L Rev 1249 [Citron 2008] at 1254.

2 Joshua Kroll, Joanna Huey, Solon Barocas, Edward Felten, Joel Reidenberg, David Robinson, and Harlan Yu, *Accountable Algorithms* (March 2, 2016) University of Pennsylvania Law Review, Vol. 165, 2017, Fordham Law Legal Studies Research Paper No. 2765268 [Kroll et al] at 41, online: https://ssrn.com/abstract=2765268.

3 Most analysis of government AI systems to date has focused on AI decision-making in the criminal justice system or on specific technologies, such as algorithmic risk assessments or facial recognition systems. In contrast, this paper considers how AI will affect government decisions in the civil and administrative justice systems.

4 Law Commission of Ontario, *The Rise and Fall of Algorithms in American Criminal Justice: Lessons for Canada*, (October 2020) [LCO Criminal AI Issue Paper], online: **https://www.lco-cdo.org/wp-content/uploads/2020/10/Criminal-AI-Paper-Final-Oct-28-2020.pdf**.

5 Law Commission of Ontario, *Legal Issues in Government AI Development*, (March 2021), [LCO Government AI Issue Paper], online: **https://www.lco-cdo.org/en/our-current-projects/ai-adm-and-the-justice-system/**.

6 Law Commission of Ontario, *Regulating AI: Critical Issues and Choices*, (April 2021) [Regulating AI], online: **https://www.lco-cdo.org/en/our-current-projects/ai-adm-and-the-justice-system/regulating-ai-critical-issues-and-choices/** and Law Commission of Ontario, *Comparing European and Canadian AI Regulation*, (November 2021) [LCO Canada/EU AI Issue Paper], online: **https://www.lco-cdo.org/en/our-current-projects/ai-adm-and-the-justice-system/comparing-european-and-canadian-ai-regulation/**.

7 Law Commission of Ontario, *AI Case Study: Probabilistic Genotyping DNA Tools in Canadian Crimninal Courts*, (June 2021) [LCO AI DNA Issue Paper], online: **https://www.lco-cdo.org/en/our-current-projects/ai-adm-and-the-justice-system/ai-case-study-pg/**.

8 A more detailed explanation of the various types of AI can be found in Paul W. Grimm, Maura R. Grossman, and Gordon V. Cormack, *Artificial Intelligence as Evidence*, 19 Nw. J. Tech. & Intell. Prop 9 (2021) [Artificial Intelligence as Evidence] at 5 to 24.

9 Canada, *Directive on Automated Decision-Making*, (2019) [Canada ADM Directive], online: **https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592**.

10 David Freeman Engstrom & Daniel E. Ho, *Algorithmic Accountability in the Administrative State* (2020) Yale J on Reg 800, online: papers.ssrn.com/sol3/papers.cfm?abstract_id=3965041 [Engstrom and Ho] at 807.

11 David Freeman Engstrom, Daniel E. Ho, Catherine M. Sharkey and Mariano-Florentino Cuéllar, *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies* (February 1, 2020) [US Federal Administrative Agencies], online: **https://ssrn.com/abstract=3551505**, at 6.

12 *Ibid*.

13 Engstrom and Ho at 813-814.

14 See generally, Cary Coglianese and David Lehr, "Transparency and Algorithmic Governance" (2019) 71:1 Admin LR 1 [Coglianese and Lehr].

15 See, for example, George Atalla, "How AI Can be a Force For Good in Government" (28 February 2019), online: *EY Consulting* **https://www.ey.com/en_gl/consulting/how-ai-can-be-a-force-for-good-in-government**; Harvard Kennedy School Ash Center, "Ten Ways Data can Make Government Better", online: *Harvard.edu* **https://datasmart.ash.harvard.edu/news/article/ten-great-ways-data-can-make-government-better-1041** and Coglianese and Lehr at 49-54.

16 Sarah Picard-Fritshe et al, *Beyond the Algorithm: Pretrial Reform, Risk Assessment, and Racial Fairness*, Center on Court Innovation, (July 2019) [Pitcard-Fritshe], online: **https://www.courtinnovation.org/sites/default/files/media/document/2019/Beyond_The_Algorithm.pdf** at 3.

17 See, for example, Australian Human Rights Commission, *Human Rights and Technology Discussion Paper,* (December 2019) [Australian Human Rights and Technology], online: **https://tech.humanrights.gov.au/sites/default/files/2019-12/TechRights_2019_DiscussionPaper.pdf** at 23. This report also discusses the risks of AI to human rights extensively.

18 This term comes from the title of a book by Cathy O'Neil, a former Wall Street data scientist and mathematician. Her 2016 book, *Weapons of Math Destruction*, popularized the idea that AI, algorithms and big data reinforce and worsen bias and discrimination in public and private sector decision-making.

19 Human Rights Watch, *Not In It For Justice,* (April 2017) [Human Rights Watch], online: **https://www.hrw.org/report/2017/04/11/not-it-justice/how-californias-pretrial-detention-and-bail-system-unfairly** at 8.

20 See generally the report by Philip Alston, the United Nations Special Rapporteur on extreme poverty and human rights discussing the implications of the digital welfare state, *Report of the Special rapporteur on extreme poverty and human rights*, 11 October 2019, 74th session, [UN Special Rapporteur], online: **https://digitallibrary.un.org/record/3834146?ln=en** and Australian Human Rights and Technology at 24.

21 US Federal Administrative Agencies at 6.

22 *Ibid* at 39-40.

23 AI Now, "Automated Decision Systems: Examples of Government Use Cases" (2019), online: **https://drive.google.com/file/d/1k4HKErSLqEWnjNOWaXg6iM9mAvcf-fhf/view** [AINow Use Cases].

24 *Ibid*.

25 US Federal Administrative Agencies at 60.

26 *Ibid* at 61.

27 *Ibid* at 32-33.

28 *Ibid* at 31.

29 AINow Use Cases at 2.

30 *Ibid*.

31 US Federal Administrative Agencies at 55-56.

32 *Ibid* at 58.

33 AINow Use Cases at 2.

34 *Ibid* at 3.

35 *Ibid* at 4.

36 *Ibid*.

37 *Ibid.*

38 *Ibid* at 5, 7.

39 AINow Use Cases at 23.

40 *Ibid.*

41 *Ibid* at 24.

42 *Ibid* at 27.

43 There are several international surveys of government use of AI and ADM by governments, including AINow Institute, *Algorithmic Accountability Policy Toolkit*, (October 2018) [AINow Accountability Toolkit], online: **https://ainowinstitute.org/aap-toolkit.pdf**; US Federal Administrative Agencies; Michele Gilman, *Poverty Algorithms: A Poverty Lawyer's Guide to Fighting Automated Decision-Making Harms in Low-Income Communities,* Data and Society (September 2020), online: **https://datasociety.net/wp-content/uploads/2020/09/Poverty-Lawgorithms-20200915.pdf**; Australian Human Rights and Technology; and *Government Use of Artificial Intelligence in New Zealand: Final Report on Phase 1 of the NZ Law Foundation's AI and Law in NZ Project*, (Wellington, 2019) [NZ AI and Law], online: **https://www.cs.otago.ac.nz/research/ai/AI-Law/NZLF%20report.pdf**._

44 Coglianese and Lehr at 8-9. Footnotes omitted.

45 AI Now Accountability Toolkit at 9.

46 See the current use surveys identified in note 45 above.

47 This situation may be changing, if slowly. Both the Government of Canada and the Province of Ontario have created "AI Registers" that identify some of the AI systems used by each respective government. As of April 2022, Ontario has disclosed a small number of AI and algorithmic systems as part of the province's Data Catalogue. For example, the Data Catalogue includes two COVID-related tools in the

"Health" section of the catalogue: (**https://data.ontario.ca/en/dataset?asset_type=algorithm&keywords_en=Health**). As of April 2022, there appeared to be four algorithmic impact assessments posted on the federal register: **https://open.canada.ca/data/en/dataset/5423054a-093c-4239-85be-fa0b36ae0b2e**.

48 *Regulating AI* at 30.

49 See generally, Darrel M. West and John R. Allen *How Artificial Intelligence is Transforming the World*, Brookings, April 24, 2018.

50 US Federal Administrative Agencies at 83.

51 See generally, the examples of Michigan's MiDAS Unemployment System (*Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold – A case study into how to automate false accusations of fraud for more than 34,000 unemployed people*, by Robert N. Charette, IEEE Spectrum, Jan, 24, 2018); and Australia's Robodebt controversy (*Robodebt: court approves $1.8bn settlement for victims of government's shameful failure*, Luke Henriques-Gomes, The Guardian, June 11, 2021); and Amnesty International "Xenophobic Machines: Discrimination Through Unregulated Use of Algorithms in the Dutch Childcare Benefits Scandal", October 2021.

52 Sandra Gabriel Mayson, *Bias In, Bias Out* (September 28, 2018), 128 Yale Law Journal 2218 (2019), University of Georgia School of Law Legal Studies Research Paper No. 2018-35 [Mayson] at 2280, online: **https://ssrn.com/abstract=3257004**.

53 US Federal Administrative Agencies at 45.

54 LCO Criminal AI Issue Paper at 11-12.

55 *Ibid* at 20-21.

56 Engstrom and Ho at 821.

57 US Federal Administrative Agencies at 76.

58 LCO Criminal AI Issue Paper at 34.

59 Lorne Sossin raises the question of who is accountable for AI in public decision making: The programmers (who may not have any input on the quality of the data, or the evolution of an AI system that continuously responds to data); the front line decision maker who relies on the output of the AI (but who may not have access to the process the AI system take to deliver its output); or the policy advisors who decided to implement an AI system? See Lorne Sossin, "Administrative Law and Artificial Intelligence" in *Litigating Artificial Intelligence*, 2021/22 eds Jill R. Presser, Jesse Beatson & Gerald Chan at 206.

60 US Federal Administrative Agencies at 83.

61 For a good example of the controversies that can result from a lack of consultation in AI policy development, see Rashida Richardson, ed., *Confronting Black Boxes: A Shadow Report of the New York City Automated Decision System Task Force*, AI Now Institute, December 4, 2019 [AINow Shadow Report], online: https://ainowinstitute.org/ads-shadowreport-2019.html.

62 In addition to the examples discussed in this paper, a small sample of additional "trustworthy AI" frameworks includes the UK Government's *Ethics, Transparency and Accountabilty Framework for Automated Decision-Making* (**https://www.gov.uk/government/publications/ethics-transparency-and-accountability-framework-for-automated-decision-making/ethics-transparency-and-accountability-framework-for-automated-decision-making**); Scotland's AI Strategy (**https://www.scotlandaistrategy.com/**); New South Wales in Australia (**https://www.digital.nsw.gov.au/policy/artificial-intelligence/ai-strategy/building-public-trust**); and New York City (**https://cities-today.com/new-york-launches-strategy-for-ethical-ai/**).

63 Surveillance Technology Oversight Project, "S.T.O.P. Condemns New York City AI Strategy for "Whitewashing Biased Tech," (October 4, 2021), online: **https://www.stopspying.org/latest-news/2021/10/14/stop-condemns-nyc-ai-strategy-for-whitewashing-biased-tech**.

64 "An EU Artificial Intelligence Act for Fundamental Rights – A Civil Society Statement," a statement of 114 European civil society organizations, (November 30, 2021) [EU AI Civil Society Statement], online: **https://www.stopspying.org/latest-news/2021/10/14/stop-condemns-nyc-ai-strategy-for-whitewashing-biased-tech at 1 and 4**.

65 European Commission, Proposal for a Regulation of the European Parliament and of The Council Laying Down Harmonised Rules on

Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, (2021) [EC AI Proposal], online: **https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206**.

66 LCO Canada/EU AI Issue Paper.

67 Ontario, "Consultation: Ontario's Trustworthy Artificial Intelligence (AI) Framework" (6 May 2021) [Ontario Trustworthy AI Framework], online: **https://www.ontario.ca/page/ontarios-trustworthy-artificial-intelligence-ai-framework-consultations**.

68 **Ontario Trustworthy AI Framework**.

69 Ontario, "Artificial Intelligence (AI) Guidance" (April 15, 2021), online: **https://www.ontario.ca/page/artificial-intelligence-ai-guidance#section-1**.

70 Law Commission of Ontario, Submission to Ontario Digital Service Consultations on Alpha Documents Concerning AI Use by Government (2020), online: **https://github.com/ongov/Transparency-Guidelines/blob/main/LCO%20Memo%20to%20ODS%20-%20June%2030%202020.pdf**.

71 Ontario Digital Service, "Action Plan – Ontario, Canada, 2021-2022", online: https://www.opengovpartnership.org/documents/action-plan-ontario-canada-2021-2022/.

72 The Government of Ontario's current inventory of AI and algorithms can be found in the Ontario Open Data Catalogue at: data.ontario.ca/group/artificial-intelligence-and-algorithms.

73 Ontario, "Beta principles for the ethical use of AI and data enhanced technologies in Ontario" (2022) [Ontario Beta Principles], online at **https://www.ontario.ca/page/beta-principles-ethical-use-ai-and-data-enhanced-technologies-ontario#:~:text=The%20Ontario%20%E2%80%9Cbeta%E2%80%9D%20principles%20complement,best%20practices%2C%20principles%20and%20frameworks**.

74 *Regulating AI* at 27-29.

75 See generally, Regulating AI 27-29 and EU AI Civil Society Statement at 1-4 among many reports discussing the need for prohibitions on very high risk AI systems.

76 See generally, *Regulating AI* at 17-47.

77 *Regulating AI* at 23-25.

78 *Ibid.*

79 LCO Criminal AI Issue Paper.

80 See generally, LCO Canada/EU AI Regulation Issue Paper for a detailed analysis of the EC AI Proposal.

81 For a good analysis of the use of AI and algorithms by child welfare agencies, see "Predictive Risk Modelling. Analytics and Child Welfare: An Examination.", an event sponsored by the NYU McSilver Institute and Fordham University Graduate School of Social Service. (November 10, 2021); online at **https://mcsilver.nyu.edu/predictive-risk-tools-in-child-welfare-practice/**. For a good analysis of AI and algorithms in policing, see Kate Robertson, Cynthia Khoo and Yolanda Song, *To Surveil and Predict, A Human Rights Analysis of Algorithmic Policing in Canada*, Citizen Lab and International Human Rights Program, University of Toronto Faculty of Law, (September 2020) [Citizen Lab], online: **https://citizenlab.ca/wp-content/uploads/2020/09/To-Surveil-and-Predict.pdf**.

82 For a sample of criticisms, see Adam Satariano, *New York Times* "British Grading Debacle Shows Pitfalls of Automated Government" (August 20, 2020), online at: **https://www.nytimes.com/2020/08/20/world/europe/uk-england-grading-algorithm.html** and Jon Porter, *The Verge* "UK ditches exam results generated by biased algorithm after student protests" (Auguat 17, 2020), online at: **https://www.theverge.com/2020/8/17/21372045/uk-a-level-results-algorithm-biased-coronavirus-covid-19-pandemic-university-applications**.

83 Allegheny Family Screening Tool (**https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx**.) For a sample of criticisms, see American Civil Liberties Union, "Family Surveillence by Algorithm" (Sept. 29, 2021), online at: **https://www.aclu.org/news/womens-rights/family-surveillance-by-algorithm-the-rapidly-spreading-tools-few-have-heard-of** and PBS Newshour, "How an Algorithm that screens for child neglect could harden racial disparities" (April 29, 2022), online at

https://www.pbs.org/newshour/nation/how-an-algorithm-that-screens-for-child-neglect-could-harden-racial-disparities.

84 This example is discussed in the section below.

85 For a sample of criticisms, see Office of the Privacy Commissioner of Canada, *Special Report to Parliament on the OPC's investigation into the RCMP's use of Clearview AI*." (June 10, 2021), online at: **https://www.priv.gc.ca/en/opc-actions-and-decisions/ar_index/202021/sr_rcmp/** and Jim Bronskill, *Global News*, "Provinces order Clearview AI to stop using facial recognition without consent." (December 14, 2021), online at: **https://globalnews.ca/news/8451440/clearview-ai-facial-recognition-order-stop/** and

86 This example is discussed extensively in the LCO's Criminal AI Issue Paper.

87 This example is discussed in the section below.

88 For a sample of criticisms, see Alejandro de la Garza, *Time Magazine,* "State's Automated Systems are Trapping Citizens in Bureaucratic Nightmares with their Lives on the Line" (May 20, 2020), online at: **https://time.com/5840609/algorithm-unemployment/** and Ryan Felton, *The Guardian,* "Michigan unemployment agency made 20,000 false fraud accusations – report" (Dec 16, 2016), online at: **https://www.theguardian.com/us-news/2016/dec/18/michigan-unemployment-agency-fraud-accusations**.

89 For a sample of criticisms, see Johana Bhuiyan, *The Guardian* "LAPD ended predictive policing amid public outcry." (NOvewmber 8, 2020), online at: **https://www.theguardian.com/us-news/2021/nov/07/lapd-predictive-policing-surveillance-reform** and Jacob Stern and Sarah Zhang, *The Atlantic,* "Do Algorithms have a Place in Policing?" (September 5, 2019), online at: **https://www.theatlantic.com/politics/archive/2019/09/do-algorithms-have-place-policing/596851/**.

90 For a sample of criticisms, see Parliament of Australia, *Design, scope, cost-benefit analysis, contracts awarded and implementation associated with the Better Management of the Social Welfare System Initiative" (June 17, 2017),* online at: **https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/SocialWelfareSystem/Report** and Luke Henriques-Gomes, *The Guardian*, "Robodebt: court approves $1.8bn settlement for victims of government's 'shameful' failure" (June 11, 2021), online at: **https://www.theguardian.com/australia-news/2021/jun/11/robodebt-court-approves-18bn-settlement-for-victims-of-governments-shameful-failure**.

91 This example is discussed in the section below.

92 **https://datajusticelab.org/about/**. Data Justice Lab at Cardiff University's School of Journalism, Media and Culture (JOMEC). More information about the Data Justice Lab's "Automating Public Services: Learning from Cancelled Systems" project is available at **https://datajusticelab.org/projects/automating-public-services-learning-from-cancelled-systems/**.

93 Joanna Redden, Jessica Brand, Ian Sander, Harry Warne, *Automated Public Services: Learning from Cancelled Systems*, Carnegie UK Trust, (Forthcoming 2022).

94 cote.

95 Colin Lecher, "What Happens When an Algorithm Cuts your Health Care" (21 March 2018)[Lecher], online: *The Verge*, **https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-arkansas-cerebral-palsy**.

96 Jacob Rosenberg, "DHS Rule Change Threatens Disabled Care, Jacob Rosenberg, Arkansas Times" (12 October 2017), online: *Arkansas Times* **https://arktimes.com/news/arkansas-reporter/2017/10/12/dhs-rule-change-threatens-disabled-care?oid=10388695**.

97 Kevin De Liban, counsel representing the applicants challenging this system, at an LCO invitational forum on automated decision-making in the civil and administrative justice system, December 2019. More information about this forum is at **https://www.lco-cdo.org/en/our-current-projects/ai-adm-and-the-justice-system/ai-and-adm-in-the-civil-administrative-justice-system/**.

98 See Lecher, above, and Catherine Rampell, "The Threat Trump Poses that gets Almost No Attention", (3 July 2017), online: *Washington Post* **https://www.washingtonpost.com/opinions/the-threat-trump-poses-that-gets-almost-no-attention/2017/07/03/151908f8-602d-11e7-8adc-fea80e32bf47_story.html**. See Rosenberg, Arkansas Times, above.

99 Karen Hao, "The coming war on the hidden algorithms that trap people in poverty" (2 December 2020), online: *MIT Technology Review* **https://www.technologyreview.com/2020/12/04/1013068/algorithms-create-a-poverty-trap-lawyers-fight-back/**.

100 See Rosenberg, *Arkansas Times*, October 12, 2017; *and Seven Individuals with Disabilities Continue the Legal Fight Against Secretive Medicaid Home Care Cuts*, Kevin De Liban, Legal Aid of Arkansas. **https://arlegalaid.org/news-events/newsroom.html/article/2017/01/30/seven-individuals-with-disabilities-continue-the-legal-fight-against-secretive-medicaid-home-care-cuts**.

101 Letcher.

102 Rashida Richardson, Jason M. Schultz & Vincent M. Southerland, "Litigating Algorithms: Challenging Government Use of Algorithmic Decision Systems" (September 2019), online: https://ainowinstitute.org/litigatingalgorithms-2019-us.pdf at 9.

103 Kevin De Liban, presentation at Law Commission of Ontario invitational forum on automated decision-making in the civil and administrative justice system, December 2019.

104 See Lecter, above: "Arkansas has said the previous, human-based system was ripe for favoritism and arbitrary decisions. 'We knew there would be changes for some individuals because, again, this assessment is much more objective,' a spokesperson told the Arkansas Times after the system was implemented. Aid recipients have pointed to a lack of evidence showing such bias in the state."

105 The Center for Democracy & Technology, "Challenging the Use of Algorithm-Driven Decision-Making in Benefits Determinations Affecting People with Disabilities" (2020), online: **https://cdt.org/wp-content/uploads/2020/10/2020-10-21-Challenging-the-Use-of-Algorithm-driven-Decision-making-in-Benefits-Determinations-Affecting-People-with-Disabilities.pdf at 11**.

106 The action was brought on behalf of 7 applicants.

107 Rashida Richardson, Jason M. Schultz, Vincent M. Southerland, AI Now Institute, *Litigating Algorithms 2019 US Report*: *New Challenges to Government Use of Algorithmic Decision Systems* at p7-8.

108 Audrey Amrein-Beardsley, "The Education Value-Added Assessment System (EVAAS) on Trial: A Precedent-Setting Lawsuit with Implications for Policy and Practice" (Spring 2019) eJournal of Educational Policy and Evaluation, online: *eric* **https://files.eric.ed.gov/fulltext/EJ1234497.pdf** at 2.

109 *Houston Federation of Teachers, Local 2415, et al. v Houston Independent School District*, 251 F Supp (3d) 1168 (SD Tex 2017) [*Local 2415*].

110 Martha Owen, counsel for the Teachers Union, at a presentation at Law Commission of Ontario invitational forum on automated decision-making in the civil and administrative justice system, December 2019.

111 *Local 2415* at 19.

112 *Ibid* at 12-13.

113 *Ibid* at 16.

114 *Ibid* at 9-10.

115 *Ibid* at 4.

116 Martha Owen presentation at Law Commission of Ontario invitational forum on automated decision-making in the civil and administrative justice system, December 2019

117 Local 2415 at 19, quoting from P. Ex 25.

118 *Ibid* at 22.

119 *Ibid* at 23 quoting City of Chicago v. Morales, 527 U.S. 41, 56 (1999).

120 U.S. Const. amend XIV, § 2.

121 Local 2415 v. HSIB ag pg. 24 quoting Gibson v. Texas Dep't of Ins., 700 F.3d 227, 238 (5th Cir. 2012).

122 *Ibid.*

123 *Nederlands Juristen Comite Voor de Mensenrechten et al v. The State of the Netherlands*, February 5, 2020, (English translation: **https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878**) [*Nederlands Juristen Comite*].

124 Jon Henley and Robert Booth, "Welfare Surveillance System Violates Human Rights,

Dutch Court Rules", *The Guardian* (5 February 2020), online: **www.theguardian.com/technology/2020/feb/05/welfare-surveillance-system-violates-human-rights-dutch-court-rules#:~:text=Welfare%20surveillance%20system%20violates%20human%20rights%2C%20Dutch%20court%20rules,-This%20article%20is&text=A%20Dutch%20court%20has%20ordered,resonate%20well%20beyond%20the%20Netherlands** and AlgorithmWatch, *Automating Society Report 2020* (October 2020) [AlgorithmWatch] at 161, online: **automatingsociety.algorithmwatch.org**.

125 AlgorithmWatch at 161, Koen Vervloesem, *How Dutch activists got an invasive fraud detection algorithm banned* (6 April 2020), online: **algorithmwatch.org/en/syri-netherlands-algorithm/**.

126 Philip Alston, "Brief by the United Nations Special Rapporteur on extreme poverty and human rights as *Amicus Curiae* in the case of NJCM c.s./De Staat de Nederlanden (SyRi) before the District Court of The Hague (case number: C/09/550982/HA ZA 18/388" (26 September 2019), online: **https://www.ohchr.org/Documents/Issues/Poverty/Amicusfinalversionsigned.pdf** [Alston Report].

127 AlgorithmWatch at 162.

128 UN Special Rapporteur at para 24.

129 *Ibid* at para 8.

130 *Nederlands Juristen Comite* at 6.49: "The court finds that it is unable to assess the correctness of the position of the State of the precise nature of SyRI because the State has not disclosed the risk model and the indicators of which the risk model is composed or may be composed. In these proceedings the State has also not provided the court with objectively verifiable information to enable the court to assess the viewpoint of the State on the nature of SyRI. The reason the State gives for this is that citizens could then adjust their conduct accordingly. This is a deliberate choice of the State. That choice also coincides with the starting point of the legislator regarding the provision of information on SyRI. The SyRI legislation does not show how the decision model of SyRI functions and which indicators are or can be used in a SyRI project (see 4.23 above for the terms decision model and indicators), i.e. which factual data make or can make the presence of a certain situation plausible."

131 See generally Vervloesem; see also AlgorithmWatch at 8.

132 Amnesty International "Xenophobic Machines: Discrimination Through Unregulated Use of Algorithms in the Dutch Childcare Benefits Scandal", (October 2021), online at: **https://www.amnesty.org/en/documents/eur35/4686/2021/en/** [Amnesty Report].

133 Niels Markus en Bart Zuidervaart, 15 Januari 2021 "Kabinet-Rutt III valt over de toeslagenaffaire", Trouw. **https://www.trouw.nl/politiek/kabinet-rutte-iii-valt-over-de-toeslagenaffaire~b17d6eac/?referrer=https%3A%2F%2Fwww.google.com%2F**; December 2020 Report – "Unprecedent Injustice" – Parliamentary Interrogation Committee Childcare Allowance "Ongekend onrechtVerslag – Parlementaire ondervragingscommissie Kinderopvangtoeslag", November 2021 "Reflection" report by the Council of State "Lessen uit de kinderopvangtoeslagzaken Reflectierapport van de Afdeling bestuursrechtspraak van de Raad van State."

134 Amnesty Report at 5-6.

135 Amnesty Report p 5.

136 July 2020 report by Data Protection Authority "Tax Authorities / Allowances: The processing of the nationality of applicants of childcare allowances." Autoriteit Persoonsgegevens, Onderzoeksrapport | z2018-22445.

137 Amnesty Report p 6.

138 December 2020 Report – "Unprecedent Injustice" – Parliamentary Interrogation Committee Childcare Allowance "Ongekend onrecht Verslag – Parlementaire ondervragingscommissie Kinderopvangtoeslag".

139 December 2020 Report – "Unprecedent Injustice" – Parliamentary Interrogation Committee Childcare Allowance "Ongekend onrecht Verslag – Parlementaire ondervragingscommissie Kinderopvangtoeslag", November 2021 "Reflection" report by the Council of State "Lessen uit de kinderopvangtoeslagzaken Reflectierapport van de Afdeling bestuursrechtspraak van de Raad van State."

140 Niels Markus en Bart Zuidervaart, 15 Januari 2021 "Kabinet-Rutt III valt over de

toeslagenaffaire", Trouw. **https://www.trouw.nl/politiek/kabinet-rutte-iii-valt-over-de-toeslagenaffaire~b17d6eac/?referrer=https%3A%2F%2Fwww.google.com%2F**.

141 In addition to the LCO's Regulating AI paper, see generally the strategies identified in note 63 and Kroll et al, Engstrom and Ho, Coglianese and Lehr, NZ AI and the Law, AINow Accountability Toolkit, EU Civil Statement, among others.

142 Ada Lovelace Institute, AI Now Institute and Open Government Partnership. *Algorithmic Accountability for the Public Sector* [Algorithmic Accountability in the Public Sector] (2021), online at: **https://ainowinstitute.org/pages/algorithmic-accountability-for-the-public-sector-report.htm**l at 3.

143 *Ibid* at 13.

144 Regulating AI at 25-27.

145 Algorithmic Accountability in the Public Sector at 5.

146 Margot E. Kaminski & Jennifer M. Urban, *The Right to Contest AI* (2021) 121:7 *Colum L Rev* 1957 [Kaminski and Urban], online: **https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3965041**.

147 *Ibid* at 1961.

148 Office of the Privacy Commissioner of Canada, "A Regulatory Framework for AI: Recommendations for PIPEDA Reform," (November 2020), online at **https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/completed-consultations/consultation-ai/reg-fw_202011/**.

149 Canada ADM Directive, s.6.4.1.

150 Scassa at 11.

151 H.B. 1655, S.B. 5527, Establishing guidelines for government procurement and use of automated decision systems in order to protect consumers, improve transparency, and create more market predictability, (Washington 2019) [Washington State 1655], online: **https://app.leg.wa.gov/billsummary?BillNumber=1655&Chamber=House&Year=2019**.

152 Washington State 1655, s. 4(1)(c).

153 *Ibid,* s. (6).

154 *EU General Data Protection Regulation (GDPR):* Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC [GDPR], OJ 2016 L 119/1, Article 22, section 3, online at https://gdpr-info.eu/art-22-gdpr/.

155 Kaminski and Urban at 2004.

156 LCO Criminal AI Issue Paper at 36-37. Footnotes omitted.

157 *Ibid*.

158 "Tactical Challenges and Practical Considerations in Litigating AI" Presser, Beatson and Robertson at 229 of *Litigating Artificial Intelligence*, 2021/2022 ed Jill R Presser, Jesse Beatson Gerald Chan.

159 Engstrom and Ho at 824-25.

160 *Ibid* at 841.

161 *Ibid.*

162 See the list of AI and algorithms currently in use in the United States in 5.3 above.

163 *Canada (Minister of Citizenship and Immigration) v Vavilov*, 2019 SCC 65 [*Vavilov*] at para 4.

164 See generally, Australian Human Rights Commission, *Human Rights and Technology: Final Report* (2021), online: **tech.humanrights.gov.au/downloads?_ga=2.224262939.1812174511.1647717023-723935921.1647717023**.

165 Toronto Declaration at Preamble para 1.

166 As the Australian Human Rights Commission notes: "New technologies can enable wide-scale surveillance online and in the physical world, which can deter individuals from legitimately sharing their views." (Australian Human Rights Discussion Paper at page 17

167 See Teresa Scassa "The Surveillant University: Remote Proctoring, AI and Human Rights", Forthcoming 2022 *Canadian Journal of*

Comparative and Contemporary Law [Surveillant University]; and Lydia X. Z. Brown, "How Automated Test Proctoring Software Discriminates Against Disabled Students" (2020) [Test Proctoring], online: **cdt.org/insights/how-automated-test-proctoring-software-discriminates-against-disabled-students/**.

168 Alice Xiang and Daniel Ho, Partnership on AI, *From Affirmative Action to Affirmative Algorithms: The Legal Challenges Threatening Algorithmic Fairness,* (November 9, 2020), online: **https://www.partnershiponai.org/affirmativealgorithms/**.

169 The LCO's Criminal AI Issue Paper discusses bias and discrimination within the context of criminal proceedings at length. The report addressed the many ways in which an AI or ADM system can be biased and the pressing need for law reform. In the LCO's view, the issues and lessons from the criminal context are applicable to other areas of government decision-making.The Canadian Human Rights Commission has raised concerns about bias in AI, particularly in policing and border controls, in their most recent annual report "Stronger Together: the Canadian Human Rights Commission's 2021 Annual Report to Parliament. **https://2021.chrcreport.ca/algorithms-in-policing.html**

170 The perspective and knowledge of those creating the AI systems tend to be integrated into the systems themselves. The share of women in tech companies today is approximately 26%. The number of women researchers and professors in AI is less than 20%. At Microsoft 4.5% of employees are Black and 6.3% are Hispanic. In many cases, AI systems are not adapted to the specific context, culture or area in which it is used, and the design and development is rarely informed by end users or populations impacted by the outcome of the AI system. Diversity in a team designing and building an AI system is helpful. However, diversity is complex and cuts along many lines including age, race, sex, socio-economic position, employment, geography and so on. See generally Genevieve Smith & Ishita Rustagi (Berkeley Haas Centre for Equity Gender and Leadership), *Mitigating Bias in Artificial Intelligence: An Equity Fluent Leadership Playbook* (2019) at 5, 40, online: **haas.berkeley.edu/wp-content/uploads/UCB_Playbook_R10_V2_spreads2.pdf**.

171 Even if we remove the discriminatory factor such as race or sex, the correlated factors can unintentionally produce bias. For example, an algorithm used to assess suitability for a loan application could unintentionally be bias against women if it is trained to weigh factors such as specific previous employment (since there are some jobs more likely occupied by women) and gaps in employment (since women are more likely to take parental leave), or part-time employment (since women are more likely to work part-time). See Jacquelyn Burkell, *The Challenges of Algorithmic Bias* (2019) Autonomy Through Cyberjustice Technologies Working Paper, Law Society of Ontario at 5, online: **www.ajcact.org/en/publications/the-challenges-of-algorithmic-bias/** and Maura Grossman paper.

172 See generally, LCO Criminal AI Issue Paper at 20-26.

173 This phrase is taken from an article by Sandra Mayson. See generally Sandra Gabriel Mayson, *Bias In, Bias Out* (2018) 128 Yale L J 2218 [Mayson] (2019), online: **https://ssrn.com/abstract=3257004**.

174 Human Rights Watch at 8.

175 See generally, Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan & Cass Sunstein, *Discrimination in the Age of Algorithms* (2018) 10 J of L Analysis 113, online: **https://doi.org/10.1093/jla/laz001** at 113. _

176 Ontario, *Promoting Public Trust and Confidence in Ontario's Data Economy* (n.d.), online: **engage.ontario.ca engage.ontario.ca/sites/default/files/discussion_paper1_eng_final.pdf**.

177 *Ibid.*

178 *Ibid* at 6-7.

179 See generally, LCO Criminal AI Issue Paper at 20-26. Note, however, the importance of understanding the differences and distinctions between how lawyers and technologists understand bias. See generally, Alice Xiang, *Reconciling Legal and Technical Approaches to Algorithmic Bias* (13 July 2020) 88:3 Tenn L Rev, online: **https://ssrn.com/abstract=3650635**.

180 See generally, Surveillant University and Test Proctoring.

181 See generally, LCO Criminal AI Issue Paper at 20-26 and Regulating AI at 40-42.

182 For an interesting summary of these initiatives, see a recent report from the Brookings

Institution, NIcol Turner Lee, Paul Resnick and Genie Barton, Algorithmic Bias detection and mitigation: Best practices and policies to reduce consumer harms, [Algorithmic Bias Detection] (2019), online at **https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/**.

183 *Human Rights Code*, R.S.O. 1990, c.H.19.

184 *Canadian Human Rights Act*, R.S.C., 1985, c.H-6.

185 *Canadian Charter of Rights and Freedoms*, Part I of the Constitution Act, 1982.

186 *Charter* s 15(1) *("*Every individual is equal before and under the law and has the right to the equal protection and equal benefit of the law without discrimination and, in particular, without discrimination based on race, national or ethnic origin, colour, religion, sex, age or mental or physical disability*").*

187 Ontario Human Rights Commission, *Guide to Your Rights and Responsibilities under the Human Rights Code* (2013) [Guide to Your Rights], online: **www.ohrc.on.ca/en/guide-your-rights-and-responsibilities-under-human-rights-code-0**.

188 In *Vriend* v. Alberta, [1989] 1 SCR 493, the court describes "the Charter's requirement of substantive, not merely formal, equality" per Cory J at para 83. *Vriend* references the earlier case of *Andrews* v. Law Society of British Columbia [1989] 1 SCR 143, where the Court said that identical treatment could produce inequality. The Court's analysis of substantive equality has evolved over the years.

189 McIntyre J in *Andrews* v. Law Society of British Columbia [1989] 1 SCR 143 at p 165 "[t]o approach the ideal of full equality before and under the law . . . the main consideration must be the impact of the law on the individual or the group concerned."

190 *Fraser* at paras 48-60.

191 *Eldridge v. British Columbia* (AG) [1997] 3 SCR 624 at paras 60-62 states "A legal distinction need not be motivated by a desire to disadvantage an individual or group in order to violate s. 15(1). It is sufficient if the effect of the legislation is to deny someone the equal protection or benefit of the law."

192 See *Fraser v Canada (Attorney General)*, 2020 SCC 28 [*Fraser*] at para 72, referencing *Brooks v Canada Safeway Ltd*, [1989] 1 SCR 1219, [1989] 2 SCJ No 42 at p. 1248, quoting James MacPherson, "Sex Discrimination in Canada: Taking Stock at the Start of a New Decade" (1980), 1 C.H.R.R. C/7, at p. C/11).

193 *Ibid.*

194 *Attorney General of Canada v Johnstone and Canadian Human Rights Commission*, 2014 FCA 110 at para 45.

195 See Burkell and Bailey at 6, referencing *Canadian National Railway Company v Canada (Human Rights Commission)*, [1987] 1 SCR 1114 at para 1134 and *Zurich Insurance Co v Ontario (Human Rights Commission)*, [1992] 2 SCR 32 at para 24.

196 *Human Rights Code* section 29.

197 *Human Rights Code* sections 36, 37(1) and 37(2).

198 *Human Rights Code* section 44.

199 *Human Rights Code* section 45.2.

200 *Moore v British Columbia (Education)* 2012 SCC 61 [*Moore*]. See also Guide to Your Rights.

201 *Ibid* at para 33 and R.B. v. Keewatin-Patricia District School Board, 2013 HRTO 1436 at para 204.

202 *Charter*, section 1.

203 *Fraser* at para 41.

204 Ontario Human Rights Commission, *Policy on ableism and discrimination based on disability* (n.d.), online: **https://www.ohrc.on.ca/en/policy-ableism-and-discrimination-based-disability/6-forms-discrimination**.

205 For a discussion of AI and systemic discrimination, see Australian Human Rights and Technology at 45, and Krishnamurthy, Vivek *AI and Human Rights Law* at p 245 Artificial Intelligence and the Law in Canada, Florian Martin-Bariteau and Teresa Scassa, eds. Lexis Nexis 2021.

206 *Egan v. Canada*, [1995] 2 S.C.R. 513 at 551-2 (25 May 1995).

207 Ontario Human Rights Commission, "Understanding discrimination in a social context"; online: **https://www.ohrc.on.ca/en/teaching-human-rights-ontario-guide-ontario-schools/appendix-5-%E2%80%93-understanding-discrimination-social-context-%E2%80%93-%E2%80%9Csocial-construction-disadvantage%E2%80%9D**.

208 See Patricia Hughes summary of academic commentators on equality jurisprudence "the Court's self-identified efforts to establish clear interpretations, has been muddled and inconsistent. For example, it has been described as "unsettled in important and troubling ways"; "confusing, unpredictable, overly burdensome and excessively formalistic"; and "bewildering, contradictory, fractured, and despair- inducing". Supreme Court of Canada Equality Jurisrudence and "Everyday Life", 2012 58 Supreme Court Law Review 245 at page 254-255.

209 See *Fraser* paras 7, 8 and 39.

210 *Hall v. Zurn Industries Limited*, 2021 HRTO 157; *Asfaha-Negusse v. Toronto (City)*, 2019 HRTO 1650 at 13; Fraser para 48 quoting *Withler*.

211 *Hughes v. Transport Canada*, 2014 CHRT 19 at 203.

212 Pakarian v. University of Toronto, 2012 HRTO 560 at para 33 quoting Pellerin v. Conseil scolaire de destrict catholique Centre-Sud, 2011 HRTO 1777.

213 See, for example, *Levan Turner v Canada Border Services Agency*, 2020 CHRT 1. This case provides a detailed and clear summary of the federal prima facie discrimination analysis at paragraphs 42-48 and summarized at paragraph 54. The case holds that a nexus between the protected ground and adverse impact must be established, and that the evidence must establish that the inference the nexus existed is more likely than any other inference.

214 *Fraser* at para 57.

215 *Ibid.*

216 See *Ibid* at para 58.

217 *Ibid* at para 59.

218 *Ibid.*

219 *Ibid.*

220 *Ibid.*

221 *Ibid* at paras 60-61.

222 Human Rights Code.

223 Fraser paras 50 and 81.

224 See Abella's reasoning at paragraph 70 and more generally 52-84; also see how the dissenting opinions describe the analysis of the majority – at paras 178 and 244. For a detailed exploration of the challenges involved in proving that the adverse impact experienced by the individual was caused by the group based characteristic see, Collen Sheppard & Mary Louise Chabot *Obstacles to Crossing the Discrimination Threshold: Connecting Individual Exclusion to Group-Based Inequalities*: *Connecting Individual Exclusion to Group-Based Inequalities* (2018) 96:1 Can Bar Rev 1 at 13, online: **https://canlii.ca/t/29zn**.

225 See Hall v. Zurn Industries Limited, 2021 HRTO 157, at para 75*Misetich v Value Village Stores Inc*, 2016 HRTO 1229 at para 51; *Cambridge Memorial Hospital and ONA (S.M.), Re*, [2017] OLAA No 22, 130 CLAS 98 at para 66; *O'Grady v Bell Canada*, 2020 FC 535 at 53; *Providence Health Care v Dunkley*, 2016 BCSC 1383: See para 94; *Levan Turner v Canada Border Services Agency*, 2020 CHRT 1.

226 See *Hall v Zurn Industries Limited*, 2021 HRTO 157, especially at paras 74-75, Applauding the Applicant's ability to gather evidence of systemic age discrimination in the form of "a chart showing salaried employees... above and below the age of 50...", the Tribunal said: "...I acknowledge the challenges faced by applicants in supporting claims of discrimination in general, and systemic discrimination in particular" and found the evidence insufficient to establish her claim. At most, the information provides additional context for the [Applicant's] own termination, but... provides little assistance in determining whether, on a balance of probabilities, the [Applicant's] dismissal was related to her age. At para 77, the Tribunal cites *Fraser* : "...in which the Court indicated..., that evidence of statistical disparity and of broader group disadvantage may demonstrate disproportionate impact, but neither is mandatory, and claimants need not show that the impugned law affects all members of a protected group in the same way." Importantly, the Respondent also had explanations for the evidence: there were reasons other individuals over 50 had been terminated.

227 Teresa Scassa, *Supreme Court of Canada Decision has Relevance for Addressing Bias in Algorithmic Decision-Making* (15 June 2018), online (blog): **www.teresascassa.ca/index.php?option=com_k2&view=item&id=278:supreme-court-of-canada-decision-has-relevance-for-addressing-bias-in-algorithmic-decision-making&Itemid=80**.

228 In *Ewert*, the Applicant was able to establish that the risk assessment tools are less accurate when applied to Indigenous inmates than when applied to non-Indigenous inmates. However, the Court found the evidentiary record insufficient to prove that the tools actually lead to harsher outcomes for Indigenous inmates. In other words, the impacted party was able to show that there was a distinction on an enumerated ground, but failed to show how the distinction led to a disadvantage. at para 79.

229 See generally, Joshua Kroll, Joanna Huey, Solon Barocas, Edward Felten, Joel Reidenberg, David Robinson, & Harlan Yu, *Accountable Algorithms* (2 March 2016) University of Pennsylvania Law Review, Vol. 165, 2017, Fordham Law L Studies Research Paper No. 2765268 at 41, online: **https://ssrn.com/abstract=2765268**.

230 COMPAS is an acronym for Correctional Offender Management Profiling for Alternative Sanctions. This controversy and related issues are discussed at length in the LCO's Criminal AI Issue Paper.

231 Julia Angwin et al., *Machine Bias*, PROPUBLICA (May 23, 2016) [ProPublica], online: **https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing**.

232 *Ibid.*

233 *Ibid.*

234 For a good summary of this debate, see Anne Washington, *How to Argue with an Algorithm: Lessons from the COMPAS ProPublica Debate* (February 2019), online: **https://ssrn.com/abstract=3357874**.

235 Mayson at 2230.

236 Richard Berk, Hoda Heidari, Shahin Jabbari, Michael Kearns and Aaron Roth, *Fairness in Criminal Justice Risk Assessments: The State of the Art*. Sociological Methods & Research (2017) [Berk et al] at 12-15, online: **https://www.researchgate.net/publication/315667137_Fairness_in_Criminal_Justice_Risk_Assessments_The_State_of_the_Art**.

237 Mayson at 2223.

238 Aziz Z. Huq, *Racial Equity in Algorithmic Criminal Justice* (20 June 2018) 68 Duke LJ, University of Chicago Public Law Working Paper No 663 at 1053, online: **https://ssrn.com/abstract=3144831**.

239 *Fraser* at para 59.

240 *Ibid.*

241 *Ibid.*

242 *Ibid*.

243 *British Columbia (Public Service Employee Relations Commission) v BCGSEU*, [1999] 3 SCR 3 [*Meiorin*] at para 54 (employment context); Miller (non-employment context).

244 *Abbey v Ontario (Community and Social Services),* 2016 HRTO 787 at para 87.

245 *Miller* at para 87.

246 *Charter*, s 1.

247 See Jacquelyn Burkell & Jane Bailey, *Unlawful Distinctions? Canadian Human Rights Law and Algorithmic Bias* (2016/2018) Canadian Yearbook of HR 227for an interesting discussion on the Supreme Court's decision in *Zurich Insurance Co v Ontario (Human Rights Commission)*, [1992] 2 SCR 32, where the majority of the Court decided that statistical profiling as a justification for insurance companies to charge male drivers under 25 years of age higher insurance premiums.

248 In other words, the enabling statue is constitutionally valid, but the individual applying the statute is alleged to be doing so in a way that breaches the *Charter*. In *Vavilov,* the court explained that it is important to "draw a distinction between cases in which it is alleged that the effect of an administrative decision being reviewed is to unjustifiably limit rights under the … *Charter*… (as was the case in Doré) and those in which the issue on review is whether a provision of the decision maker's enabling statute violates the *Charter*… " at para 57.

249 This analysis is referred to as the *Dore/Loyola* Framework and was laid out in a trilogy of Supreme Court cases: *Loyola High School v. Quebec (AG)*, 2015 SCC 12 [Loyola]; *Dore v. Barreau du Quebec*, 2012 SCC 12 [Dore]; and *Law Society of British Columbia v. Trinity Western University*, 2018 SCC 32 [Trinity Western].

250 See for example, Richard Stacey, "A Unified Model of Public Law: Charter Values and Reasonableness Review in Canada" [Unified Model of Public Law] (Summer 2021) 71 U Toronto LJ 338 and Carmelle Dieleman, "Accommodating Rights in Administrative Law: A Critique of the *Doré/Loyola* framework" (June 2021) 34 Can J Admin L & Prac 197.

251 See, for example, Unified Model of Public Law at 554.

252 See *Ktunaxa Nation v. British Columbia (Forests, Lands and Natural Resource Operations)*, 2017 SCC 54; *E.T. v. Hamilton-Wentworth District School Board*, 2017 ONCA 893; *Canadian Centre for Bio-Ethical Reform v. South Coast British Columbia Transportation Authority*, 2018 BCCA 344; *Gehl v. Canada (Attorney General)*, 2017 ONCA 319; *Strom v. Saskatchewan Registered Nurses Association*, 2020 SKCA 112; *Lauzon v. Justices of the Peace Review Council*, 2021 ONSC 6147.

253 Charlotte Baigent, "Undoing Doré: Judicial Resistance in Canadian Appellate Courts" (March 2020), 33 Can J Admin L & Prac 63 at 73.

254 *Slaight Communications Inc. v Davidson,* [1989] 1 SCR 1038, [1989] SCJ No 45; see also.

255 Spence Colburn, "Presumed Compliant: Charter Review of Statutory Grants of Discretion & The Presumption of Constitutionality" (Winter 2021) 79 UT Fac L Rev 1 at 1.

256 Spence Colburn, "Presumed Compliant: Charter Review of Statutory Grants of Discretion & The Presumption of Constitutionality" (Winter 2021) 79 UT Fac L Rev 1 at 25.

257 The *Canadian Human Rights Act* addresses intersectionality at s.3.1.

258 *Asfaha-Negusse v. Toronto (City)*, 2019 HRTO 1650 at para 13.

259 *Ibid* "The Tribunal recognizes that discrimination is often subtle and difficult to prove, and that applicants with multiple/intersecting social identities may be particularly vulnerable."; see also Moore v. The Estate of Lou Ferro, 2019 HRTO 1131 at para 33.

260 Engstrom and Ho at 815-23.

261 Ontario Beta Principles.

262 See for example, Ada Lovelace Institute, *Examining the Black Box – Tools for Assessing Algorithmic Systems*, 2020, **https://www.adalovelaceinstitute.org/case-study/examining-the-black-box/**; Nicol Turner Lee, Paul Resnick & Genie Barton, *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms* (22 May 2019), **www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/**; Independent High-Level Expert Group on Artificial Intelligence (European Commission), *ALTAI - The Assessment List on Trustworthy Artificial intelligence* (n.d.) **ec.europa.eu/newsroom/dae/document.cfm?doc_id=68342**; Roy Maurer, New York City to Require Bias Audits of AI-Type HR Technology, December 20, 2021, online at: **https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/new-york-city-require-bias-audits-ai-hr-technology.aspx**; AI Now, *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability*, Dillon Reisman, Jason Schultz, Kate Crawford, Meredith Whittaker, April 2018. **https://ainowinstitute.org/aiareport2018.pdf**.

263 Independent High-Level Expert Group on Artificial Intelligence (European Commission), *ALTAI - The Assessment List on Trustworthy Artificial intelligence* (n.d.) [*HLEG Assessment List*], online: **ec.europa.eu/newsroom/dae/document.cfm?doc_id=68342** at 16.

264 *Ibid* at 16-17.

265 Nicol Turner Lee, Paul Resnick & Genie Barton, *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms* (22 May 2019), online: **www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/**.

266 Canada, *Guideline on Service and Digital* (23 November 2021), s 4.5.3, online: **https://www.canada.ca/en/government/system/digital-government/guideline-service-digital.html**.

267 Canada ADM Directive, Appendix C, Impact Level Requirements, "Testing".

268 LCO Government AI Issue Paper at 9.

269 *Baker v Canada (Ministry of Citizen & Immigration)*, [1999] 2 SCR 817, [1999] FCJ No 39 at para 20, citing *Cardinal v Kent Institution*, [1985] 2 SCR 643, [1985] SCJ No 78 [*Kent Institution*] at 653. See also recently *Canadian Pacific Railway Company v Canada (Transportation Agency)*, 2021 FCA 69 [*Canadian Pacific*] at para 53.

270 See *Baker* paras 30-48; and Teresa Scassa, *Administrative Law and the Governance of Automated Decision Making* (2021) 54 UBC L Rev 251 at pg 1 and generally.

271 *Cardinal v Kent Institution*, [1985] 2 SCR 643, [1985] SCJ No 78 at 653, cited with approval in *Baker* at para 20. See also recently *Canadian Pacific* at para 53.

272 *Baker* at para 44, recently followed in *10.1 Inc v 2248951 Ontario Inc,* 2018 ONSC 381 at para 19.

273 See *Ibid* at para 21; recently followed in *Ali v Canada (Citizenship and Immigration)*, 2022 FC 442 at para 22; see also *Vavilov* at para 77 and *Dunsmuir v New Brunswick*, 2008 SCC 9 at para 79.

274 *Ibid* at para 33.

275 *Ibid* at para 38.

276 *Ibid* at para 37.

277 *Ibid* at paras 18-28, followed in *Vavilov* at para 77; see recently *Ogbolu v Canada (Citizenship and Immigration)*, 2022 FC 129 at para 20.

278 Treasury Board Directive on Automated Decision-Making, 2021,04,01 https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592; and the Algorithmic Impact Assessment tool **https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html**.

279 See Regulating AI at 17-49.

280 Teresa Scassa, *Administrative Law and the Governance of Automated Decision Making* (2021) 54 UBC L Rev 251.

281 See Regulating AI at 28, 33-35.

282 Canada ADM Directive, Appendix B, Impact Assessment Level.

283 Scassa at 21.

284 *Ibid*.

285 Canada ADM Directive, Appendix C, Impact Level Requirements, "Testing".

286 *Ibid.*

287 Scassa at 17.

288 *Ibid.*

289 *Ibid* referencing Jennifer Cobbe, "Administrative law and the machines of government: judicial review of automated public-sector decision-making", (2019) 39 Legal Studies 636, at 646.

290 *Old St. Boniface Residents Assn. Inc v Winnipeg (City),* [1990] SCJ No 137, [1990] 3 SCR 1170 at para 78, recently followed in *Know Your City Inc v The Corporation of the City of Brantford*, 2021 ONSC 154 at paras 35-36.

291 Teresa Scassa, *Administrative Law and the Governance of Automated Decision Making* (2021) 54 UBC L Rev 251 at 280.

292 Scassa at 17-18.

293 Canada ADM Directive, Appendix C, Impact Level Requirements, "Testing".

294 *Ibid*, s 6.2.3.

295 *Ibid*, Appendix C, Impact Level Requirements, "Testing".

296 Scassa at 24.

297 *Ibid*.

298 See Danielle Citron &. Ryan Calo, *The Automated Administrative State: A Crisis of Legitimacy* (2020) Boston University School of Law Working Paper, online: scholarship.law.bu.edu/faculty_scholarship/838 / ("But more importantly, the availability of new technological affordances invites an additional, important question: is the status quo even sufficient? Put simply, shouldn't the availability of better tools lead to higher standards for governance?" at 41).

299 Regulating AI at 10 and 19-20.

300 *Ibid* at 21 and 25.

301 Bitar, Omar and Deshaies, Benoit and Hall, Dawn, *3rd Review of the Treasury Board Directive on Automated Decision-Making*, April 22, 2022 - https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4087546.

302 Note that the issue of whether a provision of an administrative body's enabling statute violates the *Charter* is reviewed under correctness, whereas review of whether the effect of an administrative decision violates the Charter is reviewed under reasonableness. See para 57 of *Vavilov*.

303 Currently, this is an exhaustive list, however, para 70 of *Vavilov*, followed recently in *Entertainment Software Assoc. v Society Composers*, 2020 FCA 100 at para 16, and *Zeifmans LLP v Canada (National Revenue)*, 2021 FC 363 at para 15, suggests that the list could be expanded.

304 There is some room for interpretation as to what the wording in the statute must be in order to allow for appellate review; see *GSR Capital Group v The City of White Rock*, 2020 BCSC 489 at para 53, citing *Vavilov* at para 51, describing the need to look at the enabling statute to determine whether a court can curtail what a tribunal can do; see also *O'Shea/Oceanmount Community Association v Town of Gibsons*, 2020 BCSC 698 at para 46.

305 *Miller v College of Optometrists of Ontario*, 2020 ONSC 2573 at para 26, citing *Vavilov* at para 37; see also *Zarooben v The Workers' Compensation Board,* 2022 ABCA 50 at para 26; for a dispute that involves fact or both fact and law the standard is palpable and overriding error while questions of law only attract the standard of correctness.

306 *Entertainment Software Assoc v. Society Composers* 2020 FCA 100 suggests that the list could be expanded.

307 *College of Physicians and Surgeons v SJO*, 2020 ONSC 1047 at para 10.

308 *Bank of Montreal v Li*, 2020 FCA 22 at paras 27-28. It is notable that counsel framed the issue as waiving statutory entitlements. However, the court did not agree and limited specifically to interpretation of the Labour Code.

309 Note that a delegation of a final decision to an AI system could be an improper delegation and thus be subject to "correctness" review; see Scassa at 17, where she notes that the Canada Federal Directive does not address the issue of improper delegation and asks whether it may not be legally permissible to allow a machine to make a decision that a particular official is legally empowered to make.

310 *Vavilov* at para 95, recently followed in *Romania v Boros*, 2020 ONCA 216 at paras 19 and 29 and in *Farrier c Canada (procureur général), 2020* FCA 25 at paras 14-15, 19.

311 *Ibid* at para 99, , followed recently in *Saskatchewan Hospital North Battleford v Isaac*, 2022 SKCA 26 at paras 50 and 67 and in *The Owners, Strata Plan NW 2575 v Booth*, 2020 BCCA 153 at paras 21-22.

312 *Ibid* at para 90, recently followed in *English v Richmond (City)*, 2022 BCCA 442 at paras 59-60, 67 and 120.

313 *Ibid* at para 85, recently followed in *Abdikarim (Re)*, 2022 ONCA 255 at para 12.

314 *Ibid* at para 136; followed recently in *Masters v Claremont Development Corporation*, 2021 ONSC 3311 at paras 15-16.

315 The lack of justification for a decision can also be raised as a matter of procedural fairness where the enabling legislation or the duty of procedural fairness itself required them to be provided but where they were not, or where they were provided but were insufficient. See *Vavilov* ("[w]here the duty of procedural fairness or the legislative scheme mandates that reasons be given to the affected party but none have been given, this failure will generally require the decision to be set aside and the matter remitted to the decision maker… [Yet], where the reasons are provided but they fail to provide a transparent and intelligible justification… the decision will be unreasonable" at para 136).

316 *Ibid* at para 99; see also *Canada (Citizenship and Immigration) v Mason*, 2021 FCA 156 at para 26, citing *Vavilov* at para 83; see also *Reflection Productions v Ontario Media Dev Corp*, 2022 ONSC 64 at para 30, citing *Vavilov* at paras 86 and 87.

317 Richard Stacey, *A Unified Model of Public Law: Charter Values and Reasonableness Review in Canada* (2021) 71 U Toronto LJ 338 at 351

318 *Ibid* at 340.

319 Valérie Beaudouin et al, *Flexible and Context-Specific AI Explainability: A Multidisciplinary Approach* (2020), Operational Ethics Research Paper at 5-8, online: **papers.ssrn.com/sol3/papers.cfm?abstract_id=3559477**.

320 *Ibid* at 3.

321 See The Royal Society, *Explainable AI: the basics* (2019) Policy briefing, online: **https://ec.europa.eu/futurium/en/system/files/ged/ai-and-interpretability-policy-briefing_creative_commons.pdf** at 19: *[S]ystem developers might require technical details about how an AI system functions, while regulators might require assurance about how data is processed, and those subject to a decision might want to understand which factors led to an output that affected them. A single decision or recommendation might therefore need to be explained in multiple ways, reflecting the needs of different audiences and the issues at play in different situations.*

322 Bahador Khaleghi, *The What of Explainable AI* (2019), online: **www.elementai.com/news/2019/the-what-of-explainable-ai**: *What it means to explain a model can vary widely depending on the end user's role and level of sophistication, as well as specific constraints of given application environment. For example, a developer creating an AI model might prefer explanations that describe the inner workings of the model to facilitate debugging. On the other hand, a lay user who is examining the fairness of an AI model might prefer explanations that relate its input to its output.*

323 The words explainability, interpretability, understandability, comprehensibility and transparency are sometimes used interchangeably, and other times distinguished with different meanings; see Beaudoin et al. See also Alejandro Barredo Arrieta et al, *Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI* (26 December 2019) 58 Information Fusion 82 [Arrieta et a;] at 86-87 for an excellent summary of these terms.

324 Berkman Centre, *Accountability of AI Under the Law: The Role of Explanation* Finale Doshi-Velez, Mason Kortz, Ryan Budish, Chris Bavitz, Sam Gershman, David O'Brien, Kate Scott, Stuart Shieber, James Waldo, David Weinberger, Adrian Weller, Alexandra Wood, 22 Dec. 2019.

325 Arrieta et al at 93-94.

326 *Ibid* at 7-10.

327 Scassa at 281.

328 US Administrative Agencies at 12.

329 Engstrom and Ho at 841.

330 *Vavilov* at para 81, recently followed in *Canada (Citizenship and Immigration) v Galindo Camayo*, 2022 FCA 50 [*Galindo Camayo*] at para 49; see also *Biernacki v Alberta (Land and Property Rights Tribunal)*, 2022 ABCA 56 at para 23.

331 *Alexion Pharmaceuticals Inc v Canada (Attorney General)*, 2021 FCA 57 at para 12, citing *Vavilov* at paras 103-104; see also *Abdikarim (Re)*, 2022 ONCA 255 at para 12.

332 *Vavilov* at para 133; followed recently in *Galindo Camayo* at paras 50-51.

333 See, for example, Ashley Deeks, *The Judicial Demand for Explainable Artificial Intelligence* (2019) 119 Colum L Rev 1829, online: **https://www.semanticscholar.org/paper/The-Judicial-Demand-for-Explainable-Artificial-Deeks/48748a814d065c7cb450be330eb84e6188d8c4fd**.

334 *Valvilov* at para 94.

335 Stacey at 340.

336 *Vavilov* para 74.

337 *Vavilov* para 102.

338 Jennifer Raso, "AI and Administrative Law" in Florian Martin-Bariteau & Theresa Scassa, eds, *Artificial Intelligence and the Law in Canada* (LexisNexis: 2021) at 16.

339 In addition to the Arkansas Medicaid Benefits example at section 7.1 and Houston Teacher Evaluations at section 7.2 of this paper, there is also Robotdebt in Australia , in which the new automated fraud detection system reversed the "onus of proof" to the individual suspected of fraud. See Peter Whiteford "Robodebt was a fiasco with a cost we have yet to fully appreciated", The Conversation November 16, 2020 **https://theconversation.com/robodebt-was-a-fiasco-with-a-cost-we-have-yet-to-fully-appreciate-150169** *The Guardian* "Robodebt: court approves $1.8b settlement for victims of government's shameful failure", Luke Henriques-Gomes, June 11, 2021 and *Prygodicz v. Commonwealth of Australia* (No 2) [2021] FCA 634; and the MiDAS system in Michigan, see

IEEE Spectrum "Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold", Robert N. Charette, January 24, 2018; *Time Magazine*, "States' Automated Systems are Trapping Citizens in Bureaucratic Nightmares With Their Lives on the Line", Alejandro De La Garza, May 28, 2020.

340 See AINow, Litigating Algorithms 2019 US Report: *New Challenges to Government Use of Algorithmic Decision Systems*, Rashida Richardson, Jason M. Schultz and Vincent Southerland at pages 5-9 discussion of systems used in Arkansas, Idaho and Oregon to determine disability benefits. Although the legal action failed to address all of the issues, it did lead to some improvements. Note for example that in Arkansas – the state updated the system to allow the individual case worker to insert his/her own discretion to the final outcome of the benefits – page 7.

341 In the settlement agreement between the Teachers and the School Board in Houston, the School Board agreed to not use VAMs, including the EVAAS, to terminate teachers' contracts as long as the VAM score is "unverifiable". See Audrey Amrein-Beardsley, "The Education Value-Added Assessment System (EVAAS) on Trial: A Precedent-Setting Lawsuit with Implications for Policy and Practice" (Spring 2019) eJournal of Educational Policy and Evaluation, online: *eric* **https://files.eric.ed.gov/fulltext/EJ1234497.pdf** at 7.

342 Danielle Keats Citron, Technological Due Process (2008) 85 Wash U L Rev 1249 [Citron 2008]; Ryan Calo & Danielle Citron, The Automated Administrative State: A Crisis of Legitimacy, 70 Emory L.J. 797 (2021)

343 Danielle Keats Citron, *Open Code Governance* (2008) 1 U Chi L Forum 355, online: chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1430&context=uclf; Danielle Keats Citron & Frank Pasquale, *The Scored Society* (2014) 89 Wash L Rev 1 [Pasquale 2014]; Joshua A. Kroll et al, *Accountable Algorithms* (2017) 165:3 U Pa L Rev 633 online: scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=9570&context=penn_law_review; Robert Brauneis & Ellen P. Goodman (2018) *Algorithmic Transparency for the Smart City*, 20 Yale J L Tech 103 at 115–118, online: yjolt.org/algorithmic-transparency-smart-city#:~:text=As%20artificial%20intelligence%20and%20big,be%20accountable%20for%20its%20behavior.

344 Danielle Keats Citron, Technological Due Process (2008) 85 Wash U L Rev 1249 [Citron 2008] at 1254.

345 Engstrom and Ho at 854."

346 See generally Engstrom and Ho at 824-845 and the discussion later in this section."

347 US Federal Administrative Agencies.

348 Ibid at 44-45.

349 Citron & Calo at 6-7.

350 *Ibid* at 36 and 39.

351 Scassa at 15, referencing Flood and Dolling 2018 at 11.

352 *Ibid* at 15.

353 Australian Human Rights Commission, *Human Rights and Technology Discussion Paper*, December 2019 pg. 18.

354 ICES was formerly known as the Institute for Clinical Evaluation Services. More information about ICES is available at **https://www.ices.on.ca/About-ICES/Mission-vision-and-values**.

355 Some examples of ICES research include: *Predicting mortality among patients hospitalized for heart failure: derivation and validation of a clinical model*; *Mortality among patients admitted to hospitals on weekends as compared with weekdays*; and *Outcome of heart failure with preserved ejection fraction in a population-based study* for more examples see: **https://www.ices.on.ca/About-ICES/25th-anniversary/Top-25-in-25**._

356 *R. v. Dyment*, [1988] 2 SCR 417 at para 17.

357 Information and Privacy Commissioner Letter to Minister Lisa M. Thompson, September 17, 2019 [IPC Letter], online: **https://www.ipc.on.ca/wp-content/uploads/2019/09/2019-09-17-ltr-open-letter-to-minister-lisa-thompson_mg_re-discussion-paper.pdf**.

358 To Surveil and Predict at 74-89.

359 See Teresa Scassa "Data Ownership", CIGI Papers No. 187 – September 2018, **https://www.cigionline.org/sites/default/files/documents/Paper%20no.187_1.pdf**; and Alda Yuan,

"Derived Data: A novel privacy concern in the age of advanced biotechnology and genome sequencing", (August 15, 2018), Yale L. & Pol'y Rev. Inter Alia.

360 UN Special Rapporteur at 3.

361 **Netherlands**: See discussion above in section 7.3 on Benefits Fraud in the Netherlands including SyRi and Childcare Benefit fraud; **Australia**: Australia Robodebt scandal resulted in a 1.8B class action settlement - *Prygodicz v. Commonwealth of Australia* (No 2) [2021] FCA 634. See also, *The Guardian* "Robodebt: court approves $1.8b settlement for victims of government's shameful failure", Luke Henriques-Gomes, June 11, 2021 where the court noted "'heart-wrenching' stories of pain and anguish from victims". Peter Whiteford "Robodebt was a fiasco with a cost we have yet to fully appreciated", The Conversation November 16, 2020. **https://theconversation.com/robodebt-was-a-fiasco-with-a-cost-we-have-yet-to-fully-appreciate-150169** Robodebt continues to be a significant issue as opposition parties are calling for a Robodebt Royal Commission (*The Courier*, "Labor promises a Robodebt Royal Commission" by Dominic Giannini, April 30, 2022), **MiDAS Michigan** *IEEE Spectrum* "Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold", Robert N. Charette, January 24, 2018; *Time Magazine*, "States' Automated Systems are Trapping Citizens in Bureaucratic Nightmares With Their Lives on the Line", Alejandro De La Garza, May 28, 2020; Michigan is still struggling with the system which continues to cause problems see *Detroit Free Press* "Whitmer looks to replace state's unemployment system after issues with fraud, timeliness", Adrienne Roberts, March 3, 2022.

362 See section 7.3 above for a discussion of both the Netherlands welfare benefits and childcare benefits fraud.

363 R.S.O. 1990, c. F.31. There are also privacy protections in *The Personal Health Information Protection Act*, and the *Child, Youth and Family Services Act*. The *Personal Information Protection and Electronic Documents Act* (PIPEDA) applies to private-sector organizations across Canada that collect, use or disclose personal information in the course of a commercial activity – it does not apply to government. Information held by municipal governments is governed by the *Municiple Freedom of Information and Protection of Privacy Act*, R.S.O. 1990, c. M56.

364 Ontario's Freedom of Information and Protection of Privacy Act, Mini Guide **https://www.ipc.on.ca/wp-content/uploads/Resources/provincial%20guide-e.pdf**.

365 Bill 100: *Protecting What Matters Most Act (Budget Measures), 2019* [Bill 100], online: **https://www.ola.org/en/legislative-business/bills/parliament-42/session-1/bill-100#BK33**, Schedule 31.

366 David Young, "COVID and other privacy developments," (2022) [David Young Blog], online: **http://davidyounglaw.ca/compliance-bulletins/covid-and-other-privacy-developments-ontario/**.

367 Bill 100.

368 Bill 100, Schedule 31 s.4(3) Clause 42(1)(g)(ii).

369 Teresa Scassa - **http://www.teresascassa.ca/index.php?option=com_k2&view=item&id=303:ontario-budget-bill-will-amend-public-sector-privacy-laws&Itemid=80**.

370 Bill 188*, Economic & Fiscal Update Act, 2020*, Schedule 2 h**ttps://www.ola.org/en/legislative-business/bills/parliament-42/session-1/bill-188**._

371 David Young Blog.

372 *Ibid.*

373 **https://engage.ontario.ca/en/engagement-initiatives/ontarios-data-strategy**.

374 https://www.ontario.ca/page/building-digital-ontario.

375 *Ibid.*

376 *Ibid*.

377 *Ibid.*

378 See Scientific American "Anonymous Data Won't Protect your Identify", July 23, 2019, online: **https://www.scientificamerican.com/article/anonymous-data-wont-protect-your-identity/**.

379 Teresa Scassa blog **https://www.teresascassa.ca/index.php?option=com_k2&view=item&id=323:interesting-amendments-to-ontarios-health-data-and-public-sector-privacy-laws-buried-in-omnibus-bill&Itemid=80**.

380 IPC Letter.

381 Cases brought in the Superior Court and higher courts are governed by the *Rules of Civil Procedure*, RRO 1990, c C.43 [*Rules*], online: **www.ontario.ca/laws/regulation/900194** and the *Courts of Justice Act*, RRO 1990, c C.43 [*CJA*], online: **www.ontario.ca/laws/statute/90c43**. Matters brought in a Tribunal are governed by the statutory authority of the specific Tribunal and the *Statutory Powers and Procedures Act*, RSO 1990, c S.22, online: www.ontario.ca/laws/statute/90s22 (unless it is exempt in whole or in part from the latter). Judicial Review matters are subject to the *Rules* and the *Judicial Review Procedure Act,* RSO 1990, c J.1 [*JRPA*], online: **www.ontario.ca/laws/statute/90j01**. Cases where the Crown is party are also subject to the *Crown Liability and Proceedings Act, 2019*, SO 2019, c 7, Sched 17 [*CLPA*], online: **laws-lois.justice.gc.ca/eng/acts/C-50/**.

382 Judicial Review of a decision by a Tribunal is brought to the Divisional Court and the *JRPA* applies.

383 *Rules*, r 30.02(1).

384 See *Toronto Police Services Board v Ontario (Information & Privacy Commissioner),* 2009 ONCA 20 and *Brennan Center for Justice v New York City Police Department*, 160541/2016, online: **www.brennancenter.org/sites/default/files/opinion12222017.pdf**. For two examples of how courts may weigh the interests of police departments' (and perhaps analogously, other government agencies') interests in maintaining secrecy around AI use/procurement in order to prevent individuals from 'gaming the system' (and, to a limited extent, the commercial interests of private corporations who create such AI) against the public's interest in transparency about such use and procurement.

385 *Rules*, s. 31.03(2).

386 *CLPA*, s. 19(1)2.

387 See *Local 2415*. The value added metric scores were generated by complex algorithms, employing "sophisticated software and many layers of calculations." The third party vendor responsible for the algorithm, SAS, treated these algorithms and software as trade secrets, and refused to divulge them to either the school board or the teachers. COMPAS.

388 Michael Crichton & Will Boyer, *Trade Secret Law in Canada: How Rights Holders Can Secure Justice* (25 January 2019), online: **gowlingwlg.com/en/insights-resources/articles/2019/trade-secret-enforcement-in-canada/**.

389 *CJA* s. 137(2) states "A court may order that any document filed in a civil proceeding before it be treated as confidential, sealed and not form part of the public record."

390 *CJA* s. 135(1) states: "all court hearings shall be open to the public."

391 *Donovan v Sherman Estate,* 2019 ONCA 376, at para 5.

392 *Sierra Club of Canada v Canada (Minister of Finance)*, 2002 SCC 41 ["*Sierra Club*"]. The Sierra Club test is an adaptation of the *Dagenais/Mentuck* test established by the Supreme Court in the criminal context. In *Sierra Club*, the Court adapted the test to apply in civil cases where parties are seeking to keep commercial interests off the public record. For more discussion of this see Paul Schabas, *What Happens at a Bail Hearing Anyway? The Supreme Court's Troubling Retreat from the Openness Principle in Toronto Star v. Canada*, (2011) 54 Supreme Court Law Review: Osgoode's Annual Constitutional Cases Conference, online: **core.ac.uk/download/pdf/232638841.pdf**.

393 *Sierra Club* at para 53. In *Sierra Club* the Supreme Court addresses a sealing order under the *Federal Court Rules*, SOR/98-106, online: **laws-lois.justice.gc.ca/eng/regulations/sor-98-106/**. However, the Court of Appeal of Ontario has held that the *Sierra Club* test is the proper test to apply to determine if a confidentiality order should be ordered. See *Hollinger Inc (Re)*, 2011 ONCA 579 at para 11.

394 This comment is based upon informal LCO consultations with government officials on this topic.

395 Cofone, Ignacio and Strandburg, Katherine J., *Strategic Games and Algorithmic Secrecy* (October 18, 2019). 64.4 McGill Law Journal 623 (2019), NYU Law and Economics Research Paper No. 20-08, Available at SSRN: https://ssrn.com/abstract=3440878 or http://dx.doi.org/10.2139/ssrn.3440878.

396 *CLPA* s. 19(1)1; common law crown immunity is another potential defense. See *Toronto Star Newspapers Ltd v Canada*, [2005] OJ No 5533, [2005] OTC 1112.

397 Importantly, the *CLPA* does not apply to actions against the Crown for alleged breach of contract, *Charter* breaches or judicial review.

398 See *CLPA* s11(5)(c): the definition of "policy decision" is expanded to include immunity for "the manner in which a program, project or other initiative is carried out"; "the Crown's degree of supervision or control"; and "existence of management procedures or oversight mechanisms".

399 *Ibid* ss. 11(6)(e) and (f) sections immunize the Crown in making decisions on whether and the manner in which an investigation, inspection or assessment should be conducted, and whether and the manner in which an enforcement action under an *Act* is carried out.

400 One exception is *Seelster Farms v Her Majesty the Queen and OLG*, 2020 ONSC 4013. However, this decision may not be good law since the Ontario Court of Appeal decision in *Francis v Ontario,* 2021 ONCA 197 [*Francis*].

401 *Francis* at paras 111-128.

402 *Ibid* at para 128: the Court found that the *CLPA* did not bar the plaintiffs claim of systemic negligence in the solitary confinement in the prison system. The decision to use segregation in correctional facilities was a policy matter for government to determine. However, how the policy is actually applied, its process at the ground level, is operational and not closed.

403 *Deeks* at 1832.

404 See: *R v Mohan*, [1994] 2 SCR 9, [1994] SCJ No 36; *R v Abbey*, 2009 ONCA 624; *R v Sekhorn*, 2014 SCC 15; *Moore v Getahun*, 2015 ONCA 55; *Westerhof v Gee Estate*, 2015 ONCA 206; and *White Burgess Langille Inman v Abbott and Haliburton Co.* 2015 SCC 23 [*White Burgess*]. See also Todd L. Archibald, *Examining the Reliability of Expert Soft Science Evidence in the Courtroom: The Art and Science of Persuasion – Chapter IV* in *Annual Review of Civil Litigation 2014*.

405 *White Burgess* at para 19. See Gerald Chan and Mabel Lai, *Evidence: AI in the Courtroom*, in Jill Presser,Jesse Beatson & Gerald Chan, eds, *Litigating Artificial Intelligence*, (Toronto: Emond Publishing, 2021) at 275 for a discussion of these issues.

406 History has shown that expert testimony has "played an important role in miscarriages of justice." (see Honourable F. Kaufman, *The Commission on Proceedings Involving Guy Paul Morin Report* (1998), and Honourable Stephen Goudge, *Inquiry into Pediatric Forensic Pathology in Ontario* (Queen's Printer for Ontario:2008) – as quoted in John A. Olah, *A Road Map to the Admissibility of Expert Evidence: White Burgess Langille Inman v. Abbott and Haliburton Co*., online: **www.beardwinter.com/content/uploads/1451925100ARoadMaptotheAdmissibiltyofExpertEvidence.pdf)**.

407 John A. Olah, *A Road Map to the Admissibility of Expert Evidence: White Burgess Langille Inman v. Abbott and Haliburton Co*., online: **www.beardwinter.com/content/uploads/1451925100ARoadMaptotheAdmissibiltyofExpertEvidence.pdf**.

408 Insert list of examples of use of AI as evidence in criminal justice system. For example, probabilistic genotyping uses artificial intelligence algorithms to analyze DNA samples collected in police investigations or criminal prosecutions. See LCO Commissioned Report, *AI Case Study: Probabilistic Genotyping DNA Tools in Canadian Criminal Courts*, June 2021.

409 *Artificial Intelligence as Evidence* supra note 8.

410 Chan and Lai at 324.