#49806743v1

mccarthy tetrault

Unraveling the complex tapestry of AI: what's keeping organizations and legal up at night?

Webinar for Digital

Barry Sookman, McCarthy Tetrault

personal blog @ barrysookman.com

mccarthy tetrault

California Senate Bill 896

CA State Legislature page for SB896

Summary | Sponsors | **Texts** | Votes | Researc

Introduced

Bill Title: Artificial Intelligence Accountability Act.

COMMITTEE ON ARTIFICIAL INTELLIGENCE (CAI)

DRAFT FRAMEWORK CONVENTION ON ARTIFICIAL INTELLIGENCE, HUMAN RIGHTS, DEMOCRACY AND THE RULE OF LAW

C-27 44th Parliament, 1st session
November 22, 2021, to present

An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts

Short title: Digital Charter Implementation Act, 2022

Challenges, risks, & opportunities

- Regulatory developments
  - Legal compliance & best practices
  - Impacts all lifecycles of AI systems including, development, testing, deployment e.g., transparency, accountability framework, data governance, risk management frameworks
- Customer contracting impacts
  - Consulting
  - Development, testing and validation
  - Deploy/manage/license
  - SAAS
- Supply chain impacts – subcontracting, model licensing/training
- Impacts liability – regulatory compliance (AI & privacy), warranties, disclaimers, indemnities, & limits of liability.

AIDA – AI system regulatory compliance

"*AI System* is **"**a technological system that, using a model, makes inferences in order to generate output, including predictions, recommendations or decisions."

1. "High Impact system" is **"**an artificial intelligence system of which at least one of the intended uses may reasonably be concluded to fall within a class of uses set out in the schedule." [or which is added later]

2. General-purpose systems (aka generative AI system) an "artificial intelligence system that is designed for use, or that is designed to be adapted for use, in many fields and for many purposes and activities, including fields, purposes and activities not contemplated during the system's development".

3. Machine learning models "a digital representation of patterns identified in data through the automated processing of the data using an algorithm designed to enable the recognition or replication of those patterns. "

Note EU AIA: Prohibited AI systems; High Risk systems, general purpose AI models, and GPAI models with systemic risk. US Executive Order, dual use foundation models.

Initial high impact systems - AIDA

- determinations in respect of employment, including recruitment, referral, hiring, remuneration, promotion, training, apprenticeship, transfer or termination;
- determinations of whether to provide services to an individual including the type or cost of services to be provided to an individual; or the prioritization of the services to be provided to individuals;
- to process biometric information in matters relating to the identification of an individual (excluding biometric information processed with the individual's consent to authenticate their identity); or the assessment of an individual's behaviour or state of mind;
- moderation of content on an online communications platform (e.g. search engines & social media); or the prioritization of the presentation of such content;
- health care or emergency services (excluding uses of devices under the *Food and Drugs Act)*;
- by a court or administrative body in making a determination in respect of an individual who is a party to proceedings before the court or administrative body;
- to assist a peace officer in the exercise and performance of their law enforcement powers, duties and functions.

Initial high impact systems - AIDA

Minister letter to INDU Committee, November 2023.

- "*Healthcare and emergency services by definition implicate matters of health and safety, but also carry the potential for discrimination. Systems of interest would include those that triage individuals at an emergency ward, or systems that provide health advice directly to Canadians over the Internet. Medical devices incorporating AI are already captured by the Medical Device Regulations under the Food and Drugs Act and Health Canada operates a rigorous pre-market assessment system for medical devices, as well as post-market surveillance; as a result, AI systems that are medical devices would be excluded from the initial set of classes.*"

- "While AIDA does not regulate the use of AI systems by governments, it is important to recognize that many systems used in sensitive government contexts are commercially developed and managed by private sector organizations. As a result, the Government proposes to list classes of systems that are intended for sensitive public sector use, in order to ensure that such systems have undergone appropriate risk management prior to being placed on the market, or while being managed by private sector organizations, and that public sector users have the information needed to ensure that systems are being used appropriately".

New classes of high impact systems

Can be established by regulations taking into account:

- the risk of adverse impacts on the economy or any other aspect of Canadian society and on individuals, including on individuals' health and safety and on their rights recognized in international human rights treaties to which Canada is a party;
- the severity and extent of those adverse impacts;
- the social and economic circumstances of any individuals who may experience those adverse impacts; and
- whether the uses in the class or subclass that is to be added, varied or deleted are adequately regulated under another Act of Parliament or an Act of a provincial legislature.
- *Note: AIA has more limited rights to add new high risk systems: must be listed in Annex III and pose a risk of harm to health and safety, or an adverse impact on fundamental rights, and that risk is equivalent to or greater than the risk of harm or of adverse impact posed by the high-risk AI systems already referred to in Annex III.*

What risks will likely be regulated?

- AI systems with a risk of harm. “harm” defined as physical or psychological harm to an individual; damage to an individual’s property; or economic loss to an individual.
- AI systems with a risk of biased output. “biased output” defined as “content that is generated, or a decision, recommendation or prediction that is made, by an artificial intelligence system and that adversely differentiates, directly or indirectly and without justification, in relation to an individual on one or more of the prohibited grounds of discrimination set out in section 3 of the *Canadian Human Rights Act,* or on a combination of such prohibited grounds. [Subject to certain exclusions]

Prohibited AI systems - AIA

mccarthy tétrault

- AI systems that deploys subliminal techniques or purposefully manipulative or deceptive techniques to materially distort a person's behaviour by appreciably impairing the person's ability to make an informed decision
- AI systems that exploits any of the vulnerabilities persons due to their age, disability or a specific social or economic situation, to materially distort the behaviour of that person
- Certain biometric AI systems
- Social scoring systems
- AI systems for making risk assessments to assess or predict the risk of a person to commit a criminal offence
- Facial recognition systems that scrape data from Internet or CCTV footage
- Use of AI systems to infer emotions of a person in the areas of workplace and education institutions

High risk AI systems - AIA

- AI systems part of products already regulated under EU law (for safety reasons)
- AI systems listed in Annex III, but not if they do not pose a significant risk of harm, to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making:
- No- prohibited biometric systems
- Critical infrastructure
- Education and vocational training
- Employment, workers management and access to self-employment
- Access to and enjoyment of essential private services and essential public services and benefits
- Law enforcement
- Migration and border control
- Administration of justice and democratic processes

Will you be regulated?

Canada	Make Available	Manage
High impact systems	Y	Y
GenAI system	Y	Y
Machine learning model	Y (for integration in HI system)	X

EU AIA		
	Under certain conditions, providers, distributors, importers, deployers and other third parties.	

Automated decision systems

Automated Decision Making	Transparency	Explainability
GDPR	The right "not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her."	Data subjects are to be provided "meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject" "necessary to ensure fair and transparent processing in respect of the data subject".
Law 25	If making a decision based exclusively on an automated processing of PI, but "inform the person concerned accordingly" (12.1)	If making a decision based exclusively on automated processing of PI, on request, must inform of PI used, reasons and principal factors and paramaters that led to decision, and rights of correction (12.1)
CPPA	"a general account of the organization's use of any automated decision system to make predictions, recommendations or decisions about individuals that could have a significant impact on them" (62(2)(c))	If any automated decision system has a "significant impact" on the individual, on request must provide "an explanation of the prediction, recommendation or decision, the source of the information and the reasons or principal factors that led to the prediction, recommendation, or decision" (62(2)(c))

Also, California Artificial Intelligence Accountability Act (draft) Would require prior approval before a high risk automated decision system is used in State and ongoing monitoring.

AI system transparency Obligations

- Persons should be aware they are communicating with an artificial intelligence system. (s6.1 AIDA) AIA Art. 52, Calif. AI Accountability Act., s.11549.66.)
- Synthetic content
  - AIDA, for GenAI systems, obligations to label synthetic content. (s7(1))
  - AIA, AI systems that generate deepfakes or generates or manipulates text on matters of public interest must disclose it is generated by AI.) [Will this be added to AIDA?]
- Other transparency requirements:
  - AIDA, for high impact systems, the person that manages it must (as per regs) publish a plain-language description of the system that includes how the system is being used, the types of output that it generates, the mitigation measures established to mitigate risks of harm, and any other information that may be prescribed by regs. (AIDA S11(1)(f))
  - AIA, for GPAI, draw up and make publicly available a sufficiently detailed summary about the content used for training of the general-purpose AI model, according to a template provided by the AI Office. (Art.52c) [Will this be added to AIDA?]

Obligations – High Impact Systems

Before a high-impact system is made available the person who makes it available must ensure that (in accordance with regs)

- an assessment of the adverse impacts that could result from the intended use or from any other use of the system that is reasonably foreseeable has been carried out;
- take measures to assess and mitigate any risks of harm or biased output;
- test the effectiveness of the mitigation measures;
- permit human oversight of the AI system;
- the system is performing reliably and as intended and is robust even in adverse or unusual circumstances;
- maintain a manual on the system's operations;
- records are kept showing compliance and relating to the data and processes used in developing the high-impact system. (s10(1))

Obligations – High Impact Systems

A person who manages the operations of a high-impact system must (in accordance with regs)

- ensure that the requirements of the person who makes it available are met if there are reasonable grounds to believe that they have not been accomplished;
- establish measures to identify, assess and mitigate the risks of harm or biased output that could result from the use of the system and carry out tests of the effectiveness of the mitigation measures;
- ensure that humans are overseeing the system's operations;
- establish measures allowing users to provide feedback on the system's performance;
- if there are reasonable grounds to suspect that the use of the system has resulted, in serious harm or that the mitigation measures are not effective in mitigating risks of serious harm, assess whether the use of the system did actually result in serious harm or the measures are actually not effective in mitigating those risks and, if so, cease the system's operations until additional or modified measures are put in place that will mitigate risks of serious harm and comply with notification obligations;
- Keep records demonstrating compliance. (s11(1))

Obligations – Machine Learning Models

Before a machine learning model is made available, for incorporation into a high-impact system…the person who makes it available must ensure that (in accordance with regs)

- measures respecting the data used in developing the model have been established in accordance with the regulations;
- measures to identify, assess and mitigate the risks of biased output that could result from the use of the model by a high-impact system in which the model is intended to be incorporated have been established;
- a model card has been prepared;
- must keep records showing compliance and those relating to the data and processes used in developing the machine learning model. (s9(1))

Accountability Frameworks - AIDA

A person who makes a high impact system or general-purpose system available or who manages the operations of one must establish and maintain a written accountability framework that must include:

- a description of the roles and responsibilities and reporting structure for all personnel who contribute to making the AI system available or who contribute to the management of its operations and the training and training materials they received;
- policies and procedures respecting the management of risks related to, and respecting the data used by, the system;
- procedures for persons who manage the AI system to advise the person who makes it available of any use that results in serious harm or of any mitigation measures that are not effective in mitigating risks of serious harm;
- The framework must take into account the nature and size of the business and the risks of harm or biased output that could result from the use of the AI system. (s12(1), 12(3))
- Note: AIA also requires quality management systems (Art. 17)

Governance guidance from AIA

- "The risk management system should consist of a continuous, iterative process that is planned and run throughout the entire lifecycle of a high-risk AI system. This process should be aimed at identifying and mitigating the relevant risks of artificial intelligence systems on health, safety and fundamental rights. The risk management system should be regularly reviewed and updated to ensure its continuing effectiveness, as well as justification and documentation of any significant decisions and actions taken subject to this Regulation. This process should ensure that the provider identifies risks or adverse impacts and implements mitigation measures for the known and reasonably foreseeable risks of artificial intelligence systems to the health, safety and fundamental rights in light of its intended purpose and reasonably foreseeable misuse, including the possible risks arising from the interaction between the AI system and the environment within which it operates. The risk management system should adopt the most appropriate risk management measures in the light of the state of the art in AI." AIA Recital 42a. See also Art 9.

Governance guidance from AIA

- "Requirements should apply to high-risk AI systems as regards risk management, the quality and relevance of data sets used, technical documentation and record-keeping, transparency and the provision of information to deployers, human oversight, and robustness, accuracy and cybersecurity. Those requirements are necessary to effectively mitigate the risks for health, safety and fundamental rights, and no other less trade restrictive measures are reasonably available, thus avoiding unjustified restrictions to trade." AIA Recital 43

mccarthy tétrault

Governance guidance from AIA

- "High quality data and access to high quality data plays a vital role in providing structure and in ensuring the performance of many AI systems, especially when techniques involving the training of models are used, with a view to ensure that the high-risk AI system performs as intended and safely and it does not become a source of discrimination… High quality datasets for training, validation and testing require the implementation of appropriate data governance and management practices. Datasets for training, validation and testing, including the labels, should be relevant, sufficiently representative, and to the best extent possible free of errors and complete in view of the intended purpose of the system. In order to facilitate compliance… data governance and management practices should include, in the case of personal data, transparency about the original purpose of the data collection, The datasets should also have the appropriate statistical properties, including as regards the persons or groups of persons in relation to whom the high-risk AI system is intended to be used, with specific attention to the mitigation of possible biases in the datasets, that are likely to affect the health and safety of persons, negatively impact fundamental rights or lead to discrimination prohibited under Union law, especially where data outputs influence inputs for future operations ('feedback loops')". AIA Recital 54.

mccarthy tetrault

Governance guidance from AIA

- "High- risk AI systems should be designed in a manner to enable deployers to understand how the AI system works, evaluate its functionality, and comprehend its strengths and limitations. High risk AI systems, should be accompanied by appropriate information in the form of instructions of use. Such information should include the characteristics, capabilities and limitations of performance of the AI system. These would cover information on possible known and foreseeable circumstances related to the use of the high-risk AI system, including deployer action that may influence system behaviour and performance, under which the AI system can lead to risks to health, safety, and fundamental rights, on the changes that have been pre-determined and assessed for conformity by the provider and on the relevant human oversight measures, including the measures to facilitate the interpretation of the outputs of the AI system by the deployers. Transparency, including the accompanying instructions for use, should assist deployers in the use of the system and support informed decision making by them. Among others, deployers should be in a better position to make the correct choice of the system they intend to use in the light of the obligations applicable to them, be educated about the intended and precluded uses, and use the AI system correctly and as appropriate. In order to enhance legibility and accessibility of the information included in the instructions of use, where appropriate, illustrative examples, for instance on the limitations and on the intended and

mccarthy tétrault

Governance guidance from AIA

- "Having comprehensible information on how high-risk AI systems have been developed and how they perform throughout their lifetime is essential to enable traceability of those systems, verify compliance with the requirements…as well as monitoring of their operations and post market monitoring. This requires keeping records and the availability of a technical documentation, containing information which is necessary to assess the compliance of the AI system with the relevant requirements and facilitate post market monitoring. Such information should include the general characteristics, capabilities and limitations of the system, algorithms, data, training, testing and validation processes used as well as documentation on the relevant risk management system and drawn in a clear and comprehensive form. The technical documentation should be kept up to date, appropriately throughout the lifetime of the AI system. Furthermore, high risk AI systems should technically allow for automatic recording of events (logs) over the duration of the lifetime of the system." AIA Recital 46

mccarthy tétrault

Governance guidance from AIA

- "High-risk AI systems should be designed and developed in such a way that natural persons can oversee their functioning, ensure that they are used as intended and that their impacts are addressed over the system's lifecycle. AIA Recital 48
- "High-risk AI systems should perform consistently throughout their lifecycle and meet an appropriate level of accuracy, robustness and cybersecurity, in the light of their intended purpose and in accordance with the generally acknowledged state of the art." AIA Recital 49
- "The technical robustness is a key requirement for high-risk AI systems. They should be resilient in relation to harmful or otherwise undesirable behaviour that may result from limitations within the systems or the environment in which the systems operate (e.g. errors, faults, inconsistencies, unexpected situations)." AIA Recital 50

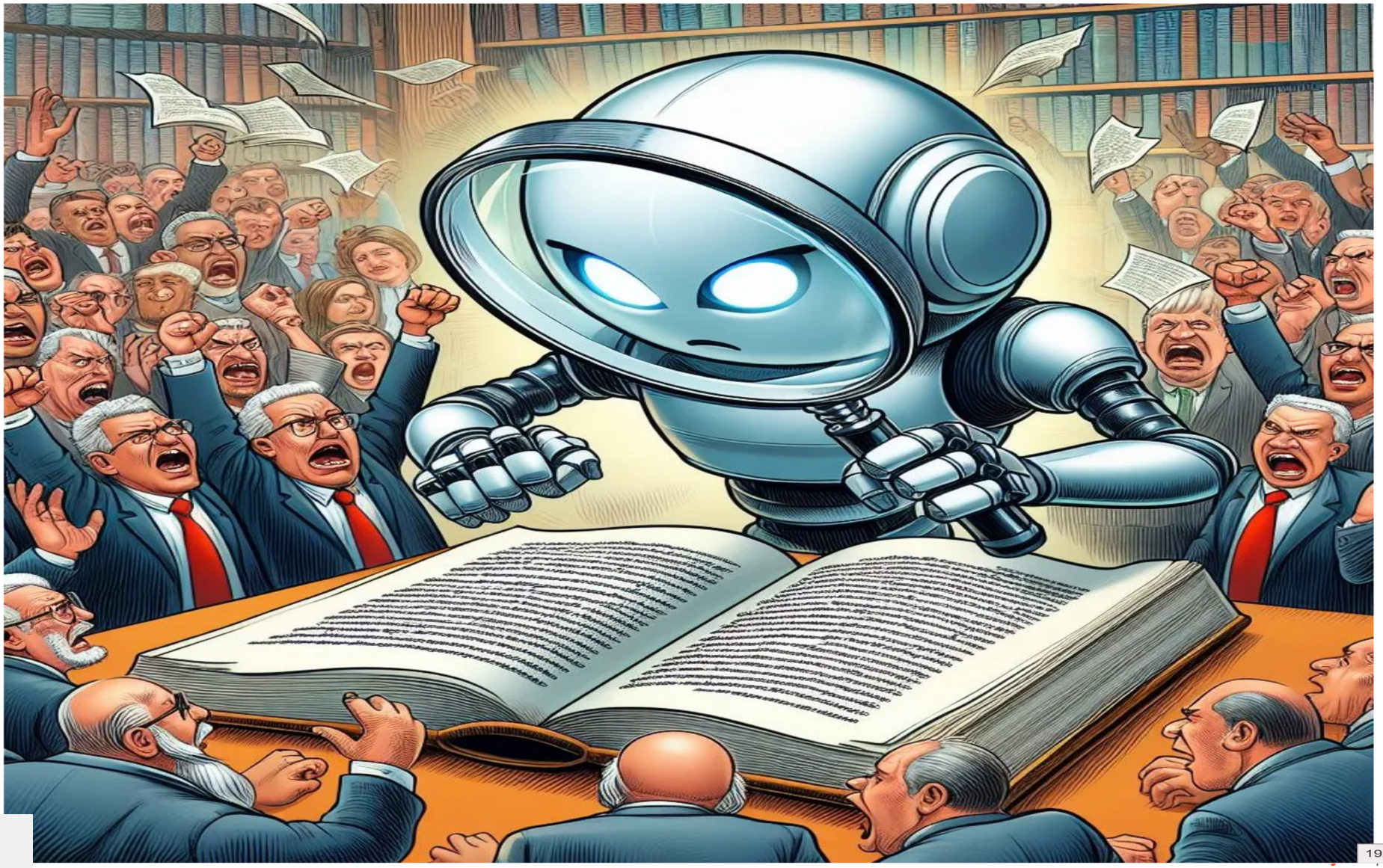
mccarthy tetrault

- Barry Sookman Resolving GenAI copyright questions: 4 court decisions

Barry Sookman BIO & EXPERTISE BOOKS & ARTICLES SPEECHES & MEDIA TERMS SUBSCRIBE

Resolving GenAI copyright infringement questions: 4 court decisions

JANUARY 3, 2024 · BARRY SOOKMAN



mccarthy tetrault

Consultation paper: Consultation on Copyright in the Age of Generative Artificial Intelligence

- Text and data mining – i.e., whether any clarification is needed on how the copyright framework applies to the use of copyright-protected works and other subject matter (e.g., a performance or sound recording) in the training of AI systems;

- Authorship and ownership of works generated by AI – i.e., how the copyright framework should apply to AI-assisted and AI-generated works; and

- Infringement and liability regarding AI – e.g., who are the persons liable when AI-generated works infringe copyright-protected works.

mccarthy tetrault

EU AIA and copyright

- "Any use of copyright protected content requires the authorization of the rightholder concer unless relevant copyright exceptions and limitations apply. Directive (EU) 2019/790 introdu exceptions and limitations allowing reproductions and extractions of works or other subject matter, for the purposes of text and data mining, under certain conditions…. Where the righ opt out has been expressly reserved in an appropriate manner, providers of general-purpos models need to obtain an authorisation from rightholders if they want to carry out text and d mining over such works." (Recital 60i)
- For GPAI, must provide a description of "information on the data used for training, testing a validation… how the data was obtained and selected".
- "For this purpose, providers of general purpose AI models should put in place a policy to respect Union law on copyright and related rights, in particular to identify and respect the reservations of rights expressed by rightholders pursuant to Article 4(3) of Directive (EU) 2019/790. Any provider placing a general purpose AI model on the EU market should comp with this obligation, regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of these general purpose AI models take place." (Recital 60j)

mccarthy tetrault

Principles for responsible, trustworthy and privacy-protective generative AI technologies

1. Legal Authority and Consent - Ensure legal authority for collecting and using personal information; when consent is the legal authority, it should be valid and meaningful.
2. Appropriate Purposes - Collection, use and disclosure of personal information should only be for appropriate purposes.
3. Necessity and proportionality - Establish the necessity and proportionality of using generative AI, and personal information within generative AI systems, to achieve intended purposes.
4. Openness - Be open and transparent about the collection, use and disclosure of personal information and the potential risks to individuals' privacy.
5. Accountability- Establish accountability for compliance with privacy legislation and principles and make AI tools explainable.
6. Individual Access - Facilitate individuals' right to access their personal information by developing procedures that enable it to be meaningfully exercised.
7. Limiting Collection, Use, and Disclosure - Limit the collection, use, and disclosure of personal information to only what is needed to fulfill the explicitly specified, appropriate identified purpose.
8. Accuracy - Personal information must be as accurate, complete, and up-to-date as is necessary for purposes for which it is to be used.
9. Safeguards - Establish safeguards to protect personal information and mitigate potential privacy risks. *See,* Principles for responsible, trustworthy and privacy-protective generative AI technologies - Office of the Privacy Commissioner of Canada

Contract considerations

- AI governance obligations including terms related risk management, data governance, recor keeping, testing and validation, data quality, bias, training, human oversight, mitigation of risks, cyber-security, robustness, monitoring and dealing with incidents, technical documentation
- Risk level determinations (for compliance) e.g. high impact, high risk, GenAI, ML models, automated decision making systems
- Compliance with laws including re bias/discrimination, safety, transparency (including under privacy and AI laws); compliance with general and AI specific laws, assistance with compliance; allocation of responsibilities
- Use of established and future standards e.g.,ISO/IEC 42001:2023: (AI Management System NIST AI Risk Management Framework, ISO/IEC CD 27090 (cybersecurity)
- Intellectual property/data/licensing issues: training data, re-use in AI models, data anonymization, ownership and use of output, confidentiality
- Liability, disclaimers and indemnities
- Assessments and compliance verification

Further readings @ barrysookman.com

- Analyzing AIDA 2.0: the problems with the proposed amendments to AIDA
- Government proposals to amend AIDA: the challenges ahead Part 2
- Minister provides proposed amendments to AIDA
- AIDA's regulation of AI in Canada: questions, criticisms and recommendations
- Proposals to amend CPPA and AIDA: the good, the bad, and the challenges ahead Part
- Legality of search engines and AI systems under PIPEDA and CPPA: Google v Privacy Commissioner
- EU AIA: agreement on Europe's new AI regulatory opus
- AIDA: my appearance before the INDU Committee
- Resolving GenAI copyright infringement questions: 4 court decisions
- Do generative AI inventions and works qualify for patents and copyrights? The Th SURYAST decisions

mccarthy tétrault

VANCOUVER

Suite 2400, 745 Thurlow Street
Vancouver BC V6E 0C5
Tel: 604-643-7100
Fax: 604-643-7900
Toll-Free: 1-877-244-7711

CALGARY

Suite 4000, 421 7th Avenue SW
Calgary AB T2P 4K9
Tel: 403-260-3500
Fax: 403-260-3501
Toll-Free: 1-877-244-7711

TORONTO

Suite 5300, TD Bank Tower
Box 48, 66 Wellington Street West
Toronto ON M5K 1E6
Tel: 416-362-1812
Fax: 416-868-0673
Toll-Free: 1-877-244-7711

MONTRÉAL

Suite 2500
1000 De La Gauchetière Street West
Montréal QC H3B 0A2
Tel: 514-397-4100
Fax: 514-875-6246
Toll-Free: 1-877-244-7711

QUÉBEC CITY

500, Grande Allée Est, 9e étage
Québec QC G1R 2J7
Tel: 418-521-3000
Fax: 418-521-3099
Toll-Free: 1-877-244-7711

NEW YORK

55 West 46th Street Suite 2804
New York NY 10036
UNITED STATES
Tel: 646-940-8970
Fax: 646-940-8972

LONDON

1 Angel Court, 18th Floor
London EC2R 7HJ
UNITED KINGDOM
Tel: +44 (0)20 7786 5700
Fax: +44 (0)20 7786 5702