



Treasury Board of Canada  
Secrétariat

Secrétariat du Conseil du Trésor  
du Canada

Canada

# 4th review of the Directive on Automated Decision-Making

## Overview and proposed modifications

Fall 2024

# Purpose

---

- Provide an overview of the 4th review of the Directive on Automated Decision-Making
- Seek feedback on policy recommendations and proposed modifications

# Background

- Providing better programs and services for Canadians sometimes involves the use of artificial intelligence (AI) to support sound decision making.
- The Government of Canada (GC) is committed to ensuring that the government's use of AI is governed with clear values, ethics and laws, and in accordance with human rights.
- Treasury Board of Canada Secretariat (TBS) is responsible for providing government-wide direction on information and data governance, information technology, security, privacy and access to information.
- The Directive on Automated Decision-Making (directive) sets rules for how federal departments can use automated systems (including AI) to make or support decisions that impact the legal rights, privileges or interests of clients. For example:
  - determining eligibility for permits and benefits
  - assessing eligibility for entry to Canada
  - deciding to hire an individual to work in the public service
  - granting market authorization for a pharmaceutical product
  - launching an investigation into an individual's conduct

# Directive on Automated Decision-Making

The directive seeks to ensure transparency, accountability and procedural fairness in the use of automated decision systems in the federal government.

It requires departments to:

- assess the impacts of automated decision systems
- be transparent
- ensure quality
- provide recourse on decisions
- report publicly on system effectiveness and efficiency

The directive is supported by the Algorithmic Impact Assessment (AIA), a questionnaire that helps departments understand and manage the risks associated with automation projects.

The directive came into effect in April 2019 and applies to systems developed or procured after April 2020. More information on the Directive and AIA can be found in the Annex.

# Reviewing the Directive

The directive must be reviewed every 2 years to keep the instrument relevant and reflect the evolving technology and regulatory landscape.

1 <sup>st</sup> review (2020-21)	2 <sup>nd</sup> review (2021-22)	3 <sup>rd</sup> review (2022-23)	4 <sup>th</sup> review (2024-25)
<ul style="list-style-type: none"> <li>✓ Strengthen transparency and quality assurance</li> <li>✓ Update references to policy instruments</li> <li>✓ Clarify requirements</li> </ul>	<ul style="list-style-type: none"> <li>✓ Author guidelines supporting interpretation of requirements</li> </ul>	<ul style="list-style-type: none"> <li>✓ Expand scope</li> <li>✓ Strengthen transparency and quality assurance</li> <li>✓ Enable inclusive approaches</li> <li>✓ Improve coherence with other policies</li> <li>✓ Assess reasons for automation</li> <li>✓ Assess impacts on persons with disabilities</li> <li>✓ Clarify requirements</li> </ul>	<p>Underway summer 2024 to 2025</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Support effective implementation</li> <li><input type="checkbox"/> Strengthen client protections</li> <li><input type="checkbox"/> Enhance assessment of impacts</li> </ul>

# Overview of key themes and issues

Three themes with 7 topics have been identified to address in the 4th review

Support effective implementation	Strengthen client protections	Enhance assessment of impacts
<ul style="list-style-type: none"> <li>• Monitor policy implementation</li> <li>• Reduce the number of organizations excluded from directive</li> <li>• Adopt internationally recognized definition of AI</li> </ul>	<ul style="list-style-type: none"> <li>• Clarify obligations and enhance impact assessment of human rights</li> <li>• Strengthen protections and assessment of impacts for persons with disabilities</li> <li>• Identify banned uses</li> </ul>	<ul style="list-style-type: none"> <li>• Clarify and enhance the AIA</li> </ul>

Slides 8 to 22 identify the recommendations and proposed updates to address the topics above. Changes indicated in bold font represent added text, and text to be removed is represented in strikethrough.

## 4<sup>th</sup> review intended outcomes

---

Implementing the proposed updates would:

- Improve protections to clients and federal institutions
- Strengthen departmental compliance with the directive
- Reinforce commitments to transparency and accountability
- Clarify and improve understanding of requirements and AIA questions to align with intent and interpretation in practice
- Reduce redundancies and misalignment across the TBS policy suite

# 1. Monitoring policy implementation

**Goal:** Increase and verify departmental compliance with the directive to support improved outcomes for clients, federal institutions and Canadian society.

Recommendations	Proposed updates
<p>Add a new reporting requirement for departments to submit a report signed by the responsible assistant deputy minister (ADM) to TBS confirming compliance with the Directive.</p>	<p><b>6.5.2 Approving a summary of how the use of the automated decision-system is fair, effective, transparent and meets the requirements of the Directive.</b>  <b>6.5.2.1 Submitting the summary to Treasury Board of Canada Secretariat prior to system production.</b></p>
<p>Add clarity to the role of the CIO of Canada related to compliance monitoring.</p> <p>Add a responsibility for TBS to publish an annual summary of the compliance reports (above) on the Open Government Portal.</p>	<p><b>8.4 Monitoring policy implementation and recommending actions to departments to improve outcomes for clients.</b></p> <p><b>8.5 Publishing an annual summary of departmental reporting requirements on the Open Government Portal.</b></p>
<p>Add a requirement for the responsible ADM to approve the completed AIA prior to its publication.</p>	<p>6.1.1 Completing, <b>approving</b>, and <del>releasing</del> <b>publishing</b> the final results of an <b>Algorithmic Impact Assessment in an accessible format on the Open Government Portal</b> prior to the production of any automated decision system.</p> <p>6.1.3 Reviewing, <del>and updating</del>, <b>and approving</b> the Algorithmic Impact Assessment on a scheduled basis, including when the functionality or scope of the automated decision system changes.</p>



## 2. Excluded organizations

**Goal:** Increase the number of organizations that are subject to the directive to expand protections and reduce risks to clients, federal institutions, and Canadian society

Recommendation	Proposed update
<p>Remove the Excluded organizations subsection 9.1.1, such that the directive would apply to Agents of Parliament</p>	<p><del>9.1.1 Agents of Parliament are excluded from this directive, including the:</del></p> <ul style="list-style-type: none"> <li><del>○ Office of the Auditor General of Canada;</del></li> <li><del>○ Office of the Chief Electoral Officer;</del></li> <li><del>○ Office of the Commissioner of Lobbying of Canada;</del></li> <li><del>○ Office of the Commissioner of Official Languages;</del></li> <li><del>○ Office of the Information Commissioner of Canada;</del></li> <li><del>○ Office of the Privacy Commissioner of Canada; and</del></li> <li><del>○ Office of the Public Sector Integrity Commissioner of Canada</del></li> </ul>

### 3. Definition of AI

**Goal:** align the GC definition of AI with a more recent and internationally-recognized definition, increasing compatibility and facilitating understanding.

Recommendations	Proposed updates
Remove the definition of AI from Appendix A of the directive	<del>Artificial intelligence - 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.</del>
Modify the definition of AI in Appendix A of the Policy on Service and Digital to align with the OECD definition	<b>An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment. (Source: <u>OECD</u>)</b>

## 4. Human rights

### Goal: Clarify obligations and enhance impact assessment of human rights

Recommendations	Proposed updates
<p>Modify the testing and monitoring requirements in the directive to more clearly reference human rights</p>	<p>6.3.1 Before <del>launching into</del> <b>a system is in</b> production, developing processes so that the data and information used by the automated decision system, as well as the system's underlying model, are tested for <b>accuracy</b>, unintended biases and other factors that may <b>unintentionally or</b> unfairly impact the outcomes <b>or infringe human rights and freedoms</b>.</p> <p>6.3.2 Developing processes to monitor the outcomes of the automated decision system to safeguard against unintentional <b>and unfair</b> outcomes and to verify compliance with <b>human rights obligations</b>, institutional and program legislation, <del>and as well as</del> this directive, on a scheduled basis.</p>
<p>Add targeted questions to the AIA to strengthen the consideration of impacts to a broader range of people</p> <p>Additional changes to the AIA related to human rights can be found on pages 3 and 6 to 9 of the <a href="#">Text Changes to the AIA tool</a> document</p>	<p>For example:</p> <ul style="list-style-type: none"> <li>• <b>Have you assessed / will you assess system performance for clients with a range of personal identity factors (for example, gender, age, race, disability, sexual orientation)?</b></li> <li>• <b>Will any of the following client groups be subject to use of the system?</b> <ul style="list-style-type: none"> <li>○ <b>Indigenous Peoples</b></li> <li>○ <b>Racialized people</b></li> <li>○ <b>Persons with disabilities</b></li> <li>○ <b>Women</b></li> <li>○ <b>2SLGBTQI+ people</b></li> <li>○ <b>Youth</b></li> <li>○ <b>Seniors</b></li> </ul> </li> <li>• <b>Describe your mitigation strategies to prevent creating or exacerbating barriers for vulnerable populations.</b></li> <li>• <b>Have you evaluated whether variables on which the system bases its decisions or recommendations could be proxies for protected characteristics?</b></li> <li>• <b>Have potential issues or harms with the existing service delivery approach been raised by clients or their representatives, the media, audits or parliamentary committees?</b></li> </ul>

## 5. Persons with disabilities

**Goal:** Strengthen protections and assessment of impacts for persons with disabilities, raising awareness of the impacts of automated decision systems

Recommendations	Proposed updates
<p>Add a requirement to document system failures and take corrective actions</p>	<p><b>6.3.3 Documenting client complaints, unexpected impacts and human overrides of the decision or assessment made by the system.</b></p> <p><b>6.3.3.1 Using findings from outcome monitoring and documented complaints, unexpected impacts and human overrides to identify issues and take corrective actions.</b></p>
<p>Add new AIA questions to capture compliance with existing accessibility standards and broaden the consideration of impacts</p> <p>Additional changes to the AIA related to protections for persons with disabilities can be found on pages 3, 4 and 6 to 8 of the <a href="#">Text Changes to the AIA tool</a> document</p>	<p>Section 3 Risk Profile:</p> <ul style="list-style-type: none"> <li>• <b>Has the system been assessed to understand whether it would create or exacerbate barriers for persons with disabilities?</b></li> </ul> <p>Section 5 About the System:</p> <ul style="list-style-type: none"> <li>• <b>Does your system comply with standard CAN/ASC EN 301 549:2024, <a href="#">Accessibility requirements for Information and Communication Technology (ICT) products and services?</a></b></li> </ul> <p>Section 8 Impact Assessment:</p> <ul style="list-style-type: none"> <li>• <b>Have you tested system performance across a diverse spectrum of disability?</b></li> </ul>

Additional protections for persons with disabilities are expected as part of the strengthening of human rights protections (slide 10) and establishment of bans (slide 13, 14)

## 6. Bans

**Goal:** Identify explicit limits or define circumstances in which automated systems pose an unacceptable risk

Recommendations	Proposed updates
Add a requirement in the Policy on Service and Digital to provide parameters for use	<b>4.4.2.4.3 Ensuring that the use of automated decision systems has clear benefits to departmental clients and Canadian society and does not pose undue risks to clients or client groups.</b>
Add a requirement and an appendix to the Directive on Service and Digital and prepare supporting guidance to state uses of AI considered unacceptable by the GC	<p>For example:  <u>Directive on Service and Digital</u>            4.4.1 The Chief Information Officer of Canada is responsible for:  <b>4.4.1.2 defining unacceptable uses of automated decision systems</b></p> <p><u>Appendix</u>  <b>AI should be used for the benefit of clients and Canadian society. The GC should not use AI that:</b></p> <ul style="list-style-type: none"> <li>• poses undue risks to clients and client groups.</li> <li>• could discriminate on protected grounds.</li> <li>• could result in undue serious negative impacts to clients or groups such as persons with disabilities, children, and vulnerable populations.</li> <li>• could negatively impact individuals' human rights.</li> </ul> <p><b>AI systems must be based on sound, scientifically validated methodologies, with a clear statistical basis for their predictions or outcomes. Consideration should be made for future potential negative impacts such as cumulative impacts of AI on individuals, groups or communities, or risks to the democracy of Canada.</b></p>

Examples of unacceptable uses are on the next slide.

## 6. Examples of unacceptable AI uses

### Unacceptable use

### Example of unacceptable use case

To manipulate or deceive in a way that alters behaviour resulting in harm or impacts to individual autonomy and fundamental freedoms

- Using deepfake technology that can spread misinformation, manipulate public opinion, and lead to increased polarization.

To score or classify people in a way that leads to unjustified censorship or surveillance or that impacts freedom of expression, privacy, and autonomy

- Collecting social behaviour and inferred, predicted or observed personal characteristics over time to create a score that impacts individuals' and groups' ability to function in society.

Biometric categorization to infer personal information about individuals such as their race, political affiliation, religion, sexual orientation, gender identify and expression, and disability status

- Using physical characteristics such as hair and eye colour to infer a person's ethnicity.

Untargeted facial recognition scraping with the intent to create or expand databases

- Taking publicly available images from the internet for identity verification

Real-time and remote biometric identification systems used in public spaces, with limited exceptions when the risk of harm is outweighed by the benefit

- Using real-time facial recognition to identify individuals that are subject to a regulatory fine.

Emotion recognition

- Relying on AI to infer emotions with the goal of determining risk, capabilities or skills or to make a decision without human oversight.

Determining the risk of a person or group committing an offence based solely on AI profiling or AI assessment of personality traits

- Taking policing action on individuals based only on the AI-predicted probability of crime score.

## 7. AIA modifications

**Goal:** Increase clarity and thoroughness of the AIA tool

Recommendations	Proposed updates
<p>Add questions in AIA sections where gaps exist or that support other areas of the 4th review</p>	<p>For example:</p> <p>Section 6 About the Algorithm:</p> <ul style="list-style-type: none"> <li>• <b>Who developed the system? Please indicate the name of the developer and solution.</b></li> <li>• <b>Describe the model being used.</b></li> </ul> <p>Section 8 Impact Assessment:</p> <ul style="list-style-type: none"> <li>• <b>Will not being subject to the use of the tool have a negative impact on an individual or groups of individuals?</b></li> <li>• <b>Are there clients or groups of clients that will be most impacted by use of the system?</b></li> <li>• <b>Does the system's effectiveness in meeting client needs exceed the potential impacts identified above?</b></li> <li>• <b>The impact of individuals being wrongfully included or overlooked by the system will be: (select from: Little to no impact; Moderate impact; High impact; Very high impact)</b></li> </ul> <p>Section 12 Procedural Fairness:</p> <ul style="list-style-type: none"> <li>• <b>Will you undertake / Have you undertaken bias mitigations in algorithm development and model testing?</b></li> <li>• <b>Does the system only consider input data and information relevant to the decision?</b></li> <li>• <b>Are the system rules directly relevant to the administrative decision?</b></li> <li>• <b>Will your team be regularly retraining the model or reviewing the rules?</b></li> </ul>
<p>Modify and add questions to respond to feedback and clarify intent</p>	<p>For example:</p> <p>Section 10 Consultations:</p> <ul style="list-style-type: none"> <li>• <b>Identify the stakeholders and indicate the project lifecycle stage in which engagement took place.</b></li> </ul> <p>Section 3 Risk profile:</p> <ul style="list-style-type: none"> <li>• <b>Does the line of business serve vulnerable populations? Are clients in this line of business particularly vulnerable?</b></li> </ul>

# Text changes to Directive and AIA

---

- Additional changes to the directive and AIA are proposed to improve clarity, reduce redundancies and align with other policy instruments.
- All proposed changes are available:
  - [Text changes to the Directive](#)
  - [Text changes to the AIA tool](#)

The following slides will indicate the change and provide a short rationale for several of the proposed changes to the Directive.



# Additional proposed changes to Directive requirements

Section	Proposed updates	Rationale
6	The Assistant Deputy Minister responsible for the program using the automated decision system, or any other <del>person</del> <b>senior official</b> named by the Deputy Head, is responsible for:	Specify that the person named by the Deputy Head should be a senior official, to ensure appropriate level of accountability
6.1.1	Completing, <b>approving</b> and <del>releasing</del> <b>publishing</b> the final results of an <b>Algorithmic Impact Assessment in an accessible format on the Open Government Portal</b> prior to the production of any automated decision system.	Avoid redundancy by combining 6.1.4 with 6.1.1
6.1.4	<del>Releasing the final results of the Algorithmic Impact Assessment in an accessible format via Government of Canada websites and any other services designated by the Treasury Board of Canada Secretariat pursuant to the Directive on Open Government.</del>	
6.2.5 6.2.5.1	<b>Obtaining and safeguarding all released versions of software components used for automated decision systems.</b> <del>All released versions of proprietary software components used for automated decision systems are delivered to, and safeguarded by, the department</del>	This requirement becomes higher level so as to apply to all software components including those developed by the GC and not only proprietary ones
6.2.6	Removal of requirement to release source code	No other GC source code is required to be released, as the previous requirements to do so have been rescinded (C.2.3.8.3, <a href="#">Archived [2020-03-31] Directive on the Management of Information Technology</a> ).
6.2.8	Documenting the decisions <b>and assessments</b> of automated decision systems in accordance with the Directive on Service and Digital, and in support of the monitoring (6.3.2), data governance (6.3.4) and reporting requirements (6.5.1)	Clarify that not only decisions, but the supporting assessments should be documented

# Additional proposed changes to the Directive requirements (continued)

Section	Proposed updates	Rationale
6.3.5	Consulting the appropriate qualified experts to review the automated decision system, <b>Algorithmic Impact Assessment and supporting documentation</b> , and publishing the complete review or a plain language summary of the findings prior to the system's production, as prescribed in Appendix C.	Clarify the documentation that must be considered for peer reviews, in alignment with the Guide to peer review of automated decision systems
6.3.7	Providing <del>adequate</del> <b>training to each</b> employee training in the design, function, and implementation <del>who develops, uses, manages, or makes decisions relating to automated decision systems on how to use, oversee, explain or maintain</del> of the automated decision system <del>to be able to review, explain, and oversee its operations,</del> as prescribed in Appendix C.	Add precision as to the type of training required across all impact levels
6.3.8	<del>Establishing strategies, plans and/or measures to support IT and business continuity management, as prescribed in Appendix C and in accordance with the Directive on Security Management</del>	Avoid redundancy with the Policy on Government Security and supporting instruments by removing the IT and business continuity management requirement
6.4.1	Providing <del>Informing</del> clients <del>with of any all applicable</del> recourse options that are available to them to challenge the administrative decision.	Editorial changes for clarity.
6.4.1.1	<b>Ensuring that recourse options are timely, effective, and easy to access.</b>	Specify parameters for the recourse options that are made available.
6.5.1	Publishing information on the effectiveness and efficiency of the automated decision system in meeting program objectives on <del>a website or service designated by the Treasury Board of Canada Secretariat</del> <b>the Open Government Portal</b>	Clearly specify that the Open Government Portal is the preferred location for publication

# Additional proposed changes to Appendix A and B

Section	Proposed updates	Rationale
7.1  7.2	<p><del>Consequences of non-compliance with this directive can include any measure allowed by the <i>Financial Administration Act</i> that the Treasury Board would determine as appropriate and acceptable in the circumstances.</del></p> <p>For an outline of the consequences of noncompliance, refer to the <i>Framework for the Management of Compliance</i>, Appendix C: Consequences for Institutions and Appendix D: Consequences for Individuals.</p>	Address redundancy with the Policy on Service and Digital. Align with approach for other Directives that support the Policy on Service and Digital.
Appendix A	<p>Add, update and remove definitions</p> <ul style="list-style-type: none"> <li>- Add: production, proprietary</li> <li>- Update: algorithmic impact assessment, automated decision system, test environment</li> <li>- Remove: artificial intelligence, procedural fairness (added to preamble), source code</li> </ul>	Reduce duplication and inconsistency across policy instruments. Remove definitions for words that are not included in the Directive. Provide consistent terminology to support instrument.
Appendix B	<p>For example (Level 1):  <b>The context in which the system is operating likely has low levels of risk associated with it. This may be because of:</b></p> <ul style="list-style-type: none"> <li>▪ <b>the identity factors of the clients that may be impacted;</b></li> <li>▪ <b>the line of business and the decision that the system is supporting;</b></li> <li>▪ <b>the type of technology being used.</b></li> </ul> <p>The decision will likely have little to no, <b>easily reversible and brief impacts on some of:</b> impact on:</p> <ul style="list-style-type: none"> <li>▪ the rights of individuals or communities;</li> <li>▪ the equality, dignity, privacy, and autonomy of individuals;</li> <li>▪ the health or well-being of individuals or communities;</li> <li>▪ the economic interests of individuals, entities, or communities;</li> <li>▪ the ongoing sustainability of an ecosystem.</li> </ul> <p>Level 1 decisions will often lead to impacts that are reversible and brief.</p> <p><b>The data used by the system likely presents low levels of risk. This may be because of:</b></p> <ul style="list-style-type: none"> <li>▪ <b>the sensitivity of the data (such as the use of non-personal or unclassified information);</b></li> <li>▪ <b>the use of structured data;</b></li> <li>▪ <b>the data collection approach.</b></li> </ul>	Expand how each impact level is described to more accurately reflect the assessment areas in the AIA.

# Additional proposed changes to Appendix C

Section	Proposed updates	Rationale
Appendix C	<p><u>Notice</u> (Level 1)  <del>None</del> <b>Plain language notice posted through all service delivery channels in use (Internet, in person, mail or telephone).</b></p> <p><u>Notice</u> (Level 3, 4)  <b>In addition, the notice must direct clients to the published explanation required under Explanation level 1.</b> <del>publish documentation on relevant websites about the automated decision system, in plain language, describing: how the components work; how it supports the administrative decision; results of any reviews or audits; and a description of the training data, or a link to the anonymized training data if this data is publicly available.</del></p>	<p>Require notice for the use of all automated decision-systems to strengthen transparency and accountability. Align with impact level 2 requirements</p> <p>Align notice and explanation requirements for consistency in communications and decreased duplication.</p>
Appendix C	<p><u>Explanation</u> (Level 1)  - Updated for plain language and to include “results of any reviews or audits”</p> <p><u>Explanation</u> (Level 3, 4)  <b>The explanation from Level I is published.</b>  <b>In addition, a more detailed, meaningful, plain language, explanation is provided to the client with any decision that results in the denial of a benefit or service, or involves a regulatory action.</b>  <b>This explanation must inform the client in plain language of the reason or justification of the administrative decision. This involves a clear and client-focused description of how the system came to the output it did, including:</b></p> <ul style="list-style-type: none"> <li>▪ <b>the principal factors that led to it, such as a description of the decision tree, scoring or weights of certain factors, and</b></li> <li>▪ <b>how the system output was used by human officers</b></li> </ul> <p><b>The client should also be provided with the link to the published level I explanation.</b></p>	<p>Clarify explanation requirements. Increase consistency and reduce duplication across explanation levels (and notice requirements).</p>

# Additional proposed changes to Appendix C (continued)

Section	Proposed updates	Rationale
Appendix C	<p><u>Peer review</u> Remove option to publish specifications of the automated decision system in a peer-reviewed journal.</p>	<p>A journal's peer review process is unlikely to be consistent with the breadth of technical and ethical assessment expected for the peer review under the Directive.</p>
Appendix C	<p><u>Gender-based Analysis Plus</u> <b>Ensure that the Gender-based Analysis Plus addresses the following issues includes:</b></p> <ul style="list-style-type: none"> <li>▪ <b>An assessment of how impacts of the automation project (including the system, data and decision) on might impact gender and/or other identity factors different population groups. This includes considering of the impacts of the system and data used in the project, as well as the likely impact of the final decision. Where possible, cite the data used to assess the impacts. It is recommended that the data be gender-disaggregated and include other intersecting identity factors such as age, intersectionality of personal attributes such as gender, disability and race. If the data is unavailable, identify where the data gaps exist;</b></li> <li>▪ <b>Details of planned or existing measures to address risks identified through the Gender-based Analysis Plus or other assessments.</b></li> </ul>	<p>Expand and clarify expectations for what the Gender-based Analysis Plus should include.</p>

## Additional proposed changes to Appendix C (continued 2)

Section	Proposed updates	Rationale
Appendix C	<p>Human-in-the-loop for decisions <b>Ensuring human involvement</b></p> <p><del>Decisions may be rendered without direct human involvement.</del> <b>The system can make decisions and assessments without direct human involvement. Humans are involved in system quality assurance and are notified when the system produces undesirable outputs and can override decisions made by the system.</b></p>	Clarify that even if the system is making the decision, humans are still involved in aspects such as quality assurance and overrides.
Appendix C	<p><del>Approval for the system to operate (Level 1 and 2)</del></p> <p><del>None</del> <b>Assistant Deputy Minister responsible for the program.</b></p>	Clarify that the ADM of the program needs to approve the system.

# Tell us what you think

- We are seeking your input on the recommendations and directive edits as part of the 4th review of the directive.
- Please respond to the survey where you will be asked to:
  - confirm if the recommendations help to achieve the topic goals
  - share specific edits to the updated text of the directive and AIA
  - identify any concerns or gaps
  - provide input on the approach to the "bans" topic
- We recommend that you refer to the following documents as you complete the survey:
  - Overview of the 4th review of the Directive on Automated Decision-Making
  - Text changes to the Directive
  - Text changes to the AIA tool
- The survey will be open from November 19, 2024 to January 8, 2025

## Next steps

- Review the 4<sup>th</sup> review materials and respond to the survey by January 8

### Timeline

Fall 2024	Winter 2025	Spring 2025	Ongoing
Consultation			
Analysis & Incorporate Feedback			
	Governance & Approvals		
		Publish	
			Support Departments



# Questions?

Reach out to the TBS  
Responsible Data and AI team at  
[ai-ia@tbs-sct.gc.ca](mailto:ai-ia@tbs-sct.gc.ca) if you have  
questions or require alternative  
ways to provide input



# Annex

# Overview of the Directive on Automated Decision-Making

## Directive on Automated Decision-Making Requirements

### Algorithmic Impact Assessment

#### *Understand*

- AIA before production
- AIA when scope or functionality changes
- Publication of AIA results

### Transparency

#### *Communicate*

- Notice before decisions
- Explanation after decisions
- Access to components
- Release of source code
- Documentation of decisions
- Publication of results in meeting program objectives

### Quality assurance

#### *Prevent*

- Testing and monitoring of outcomes
- Data quality
- Data governance
- Peer review
- GBA Plus
- Employee training
- Continuity management
- Security
- Consultation with legal services
- Human intervention

### Recourse

#### *Correct*

- Recourse options to challenge decisions

# Scope of the directive

---

- Applies to services
- Administrative decision
- Systems that make a decision or related assessment
- Simple and complex technologies (not just AI)
- Systems in production

For more information, refer to the [Guide on the Scope of the Directive](#)

# Decisions and related assessments

The directive applies to systems that make decisions and related assessments, including systems that:

- present relevant information to the decision-maker,
- alert the decision-maker of unusual conditions,
- present information from other sources (“data matching”),
- provide assessments, for example by generating scores, predictions, or classifications,
- recommend one or multiple options to the decision-maker,
- make partial or intermediate decisions as part of a decision-making process, or
- make the final decision.

**Recommendations**



**Decisions**

# Algorithmic impact assessment (AIA)

- mandatory risk assessment tool
- questionnaire determines the impact level of an automated decision-system
- composed of 51 risk and 34 mitigation questions
- assessment scores are based on many factors including systems design, algorithm, decision type, impact and data
- developed based on best practices in consultation with both internal and external stakeholders
- developed in the open, and available to the public for sharing and re-use under an open license

The screenshot shows the 'Algorithmic Impact Assessment' tool interface. At the top, it features the Government of Canada logo and the text 'Government of Canada' and 'Gouvernement du Canada'. Below this is a dark header with the title 'Algorithmic Impact Assessment'. A breadcrumb trail shows 'Home > Open Government'. The main heading is 'Algorithmic Impact Assessment'. A light blue information box states: 'Information in the AIA is only stored locally on your computer, and the Government of Canada does not have access to the information you place into the tool. If you wish to keep your work, please save the data locally for future use by using the 'Save' button. You can reload a previously saved AIA form using the 'Upload JSON File' button.' Below this are three buttons: 'Save' (green), 'Upload JSON File', and 'Start Again'. A section titled 'Navigate to a Specific Page (Out of 13)' contains a dropdown menu set to 'Section 8: Impact Assessment' and a progress bar showing 'Page 8 of 13'. The 'Impact Assessment' section contains three questions with radio button options:

- Will the system only be used to assist a decision-maker?  
 Yes  
 No
- Will the system be replacing a decision that would otherwise be made by a human?  
 Yes  
 No
- Will the system be replacing human decisions that require judgement or discretion?  
 Yes  
 No

The final question is: 'Is the system used by a different part of the organization than the ones who developed it?' with an  Yes option.

# Algorithmic Impact Assessment (AIA) areas

## Risk areas

- Risk profile
- Reasons for automation
- System
- Algorithm
- Decision
- Impact assessment
- Data source and type

## Mitigation areas

- Consultations  
(internal and external)
- Data quality
- Procedural fairness
- Privacy

# The AIA process

