Treasury Board of Canada Secretariat   Secrétariat du Conseil du Trésor du Canada

Canada

Canadian federal government's approach to Responsible AI

Digital Transformation in Government 2023

June 16, 2023

Contact: Benoit Deshaies, Director responsible data and artificial intelligence, Treasury Board of Canada Secretariat
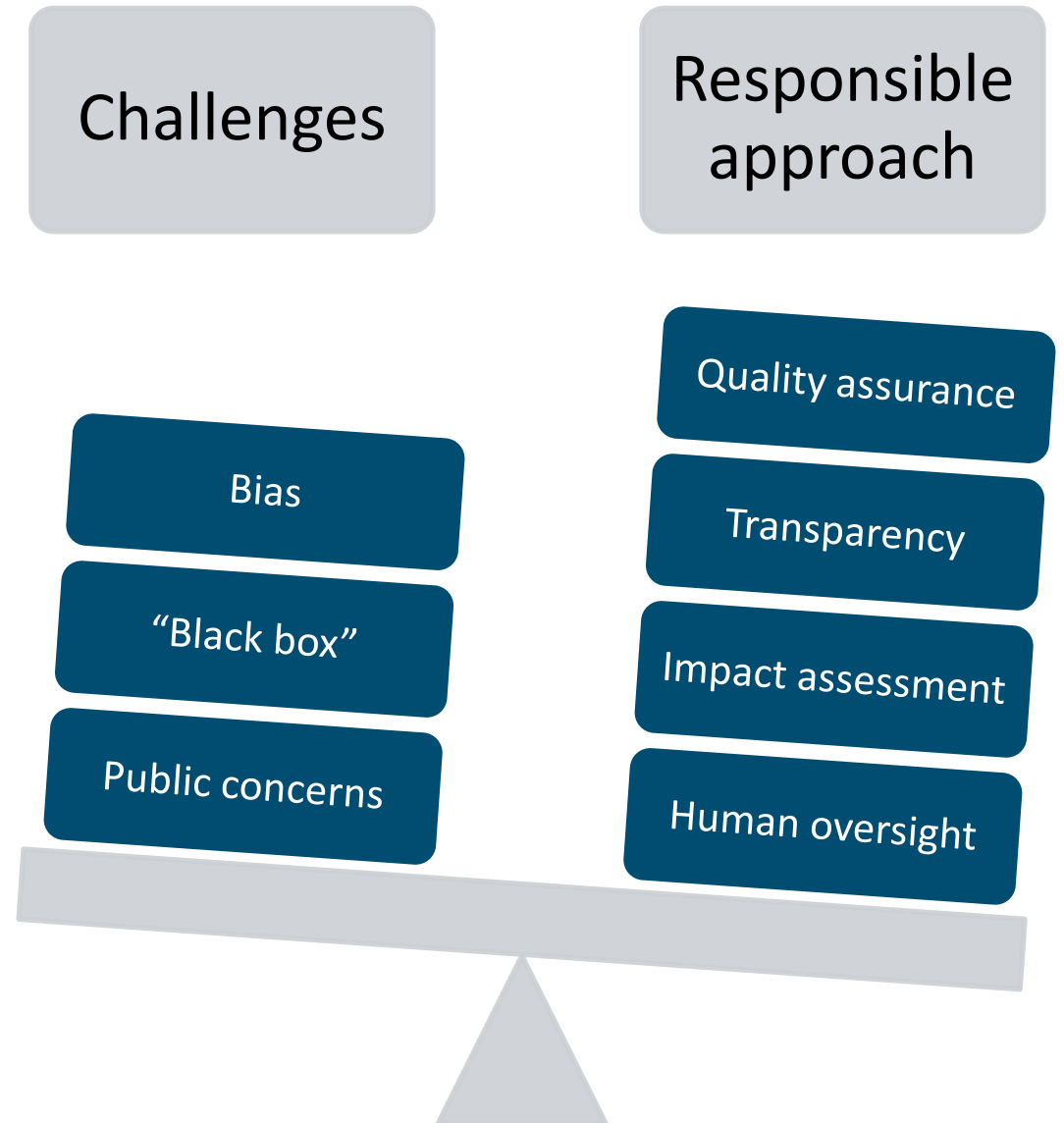
A younger version of @MrDeshaies

Automated decisions can be more

- Efficient
- Accurate
- Consistent
- Interpretable

A responsible approach will help create the conditions for success.

Challenges

Responsible approach

Quality assurance

Bias

Transparency

"Black box"

Impact assessment

Public concerns

Human oversight

Waves of algorithmic governance

Principles

Frameworks

Technical guides

- **Digital Nations** Shared Approach for the Responsible Use of AI by Governments (Nov 2018)
- **OECD** Principles on Artificial Intelligence (May 2019)

- **Canada** Directive on Automated Decision-Making (Apr 2019)
- **NIST** AI Risk Management Framework (Jan 2023)
- **France** Loi pour une République numérique

- **IEEE P7000** Standard Model Process for Addressing Ethical Concerns during System Design
- **ISO/IEC DTR 24027** Bias in AI Systems and AI Aided Decision Making

Waves of algorithmic governance (one example)

Principles	Fairness

Frameworks	• Test data and models for bias, monitor outcomes (substantive fairness) • Notice, explanation, recourse, human training and in-the-loop (procedural fairness)

Technical guides	• Fairness metrics equations (false positive parity, true positive rate, etc.) • Lifecycle approach (labelling, training, outcome testing)

Government | Gouvernement
of Canada | du Canada

Search Canada.ca

Jobs ▾ | Immigration ▾ | Travel ▾ | Business ▾ | Benefits ▾ | Health ▾ | Taxes ▾ | More services ▾

Home ➜ How government works ➜ Policies, directives, standards and guidelines

Directive on Automated Decision-Making

The Government of Canada is increasingly looking to use artificial intelligence to make or support administrative decisions to improve service delivery. The government is committed to using artificial intelligence in a manner that is compatible with core principles of administrative law such as transparency, accountability, legality, and procedural fairness. Understanding that this technology is changing rapidly, this directive will continue to evolve to ensure that it remains relevant.

Date modified: 2023-04-25

Expand all | Collapse all

▸ 1. Effective date

▸ 2. Authorities

▸ 3. Definitions

▾ 4. Objectives and expected results

4.1   The objective of this directive is to ensure that automated decision systems are deployed in a manner that reduces risks to clients, federal institutions and Canadian society, and leads to more efficient, accurate, consistent and interpretable decisions made pursuant to Canadian law.

4.2   The expected results of this directive are as follows:

     4.2.1   Decisions made by federal institutions are data-driven, responsible and comply with procedural fairness and due process requirements.

▸ More information ℹ

▸ Hierarchy ⧉

▸ Archives 🗄

🖶 Print-friendly | XML

- ✓ **People are informed about when and how automation is used;**

- ✓ **People are provided with meaningful explanations about decisions affecting them;**

- ✓ **The decisions are fair and accurate; and**

- ✓ **The potential negative impacts of automation are identified and minimized.**

Overview of the Directive on Automated Decision-Making

Directive on Automated Decision-Making Requirements

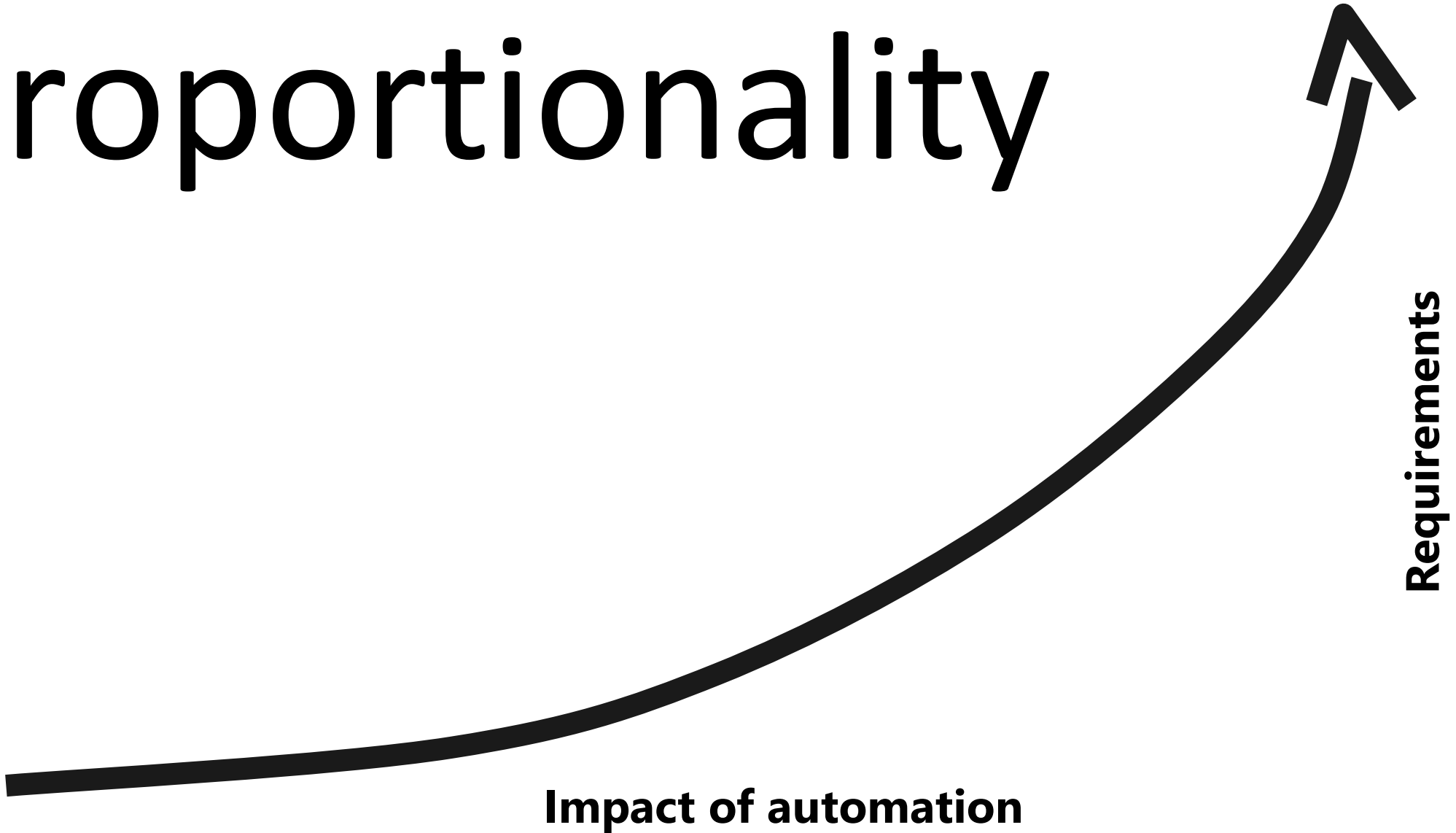| Algorithmic Impact Assessment	Transparency	Quality assurance	Recourse
Understand	*Communicate*	*Prevent*	*Correct*
▪ AIA before production ▪ AIA when scope or functionality changes ▪ Publication of AIA results	▪ Notice before decisions ▪ Explanation after decisions ▪ Access to components ▪ Release of source code ▪ Documentation of decisions ▪ Publication of results in meeting program objectives	▪ 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 options to challenge decisions

Proportionality

Requirements

Impact of automation

Why did the government AI refuse to play chess with the bureaucrat?

Because it heard that in government, every move takes forever and by the time it's your turn, the rules have changed!

- ChatGPT (GPT-4)

Principles for responsible generative AI use

F air

A ccountable

S ecure

T ransparent

E ducated

R elevant

Ensure that content from these tools does not include or amplify biases and that it complies with human rights, accessibility, and procedural and substantive fairness obligations

Accountable

Take responsibility for the content generated by these tools. This includes making sure it is factual, legal, ethical, and compliant with the terms of use

Ensure that the infrastructure and tools are appropriate for the security classification of information, and that privacy and personal information are protected

Identify content that has been produced using generative AI. Notify users that they are interacting with an AI tool. Document decisions and be able to provide explanations if tools are used to support decision-making

Learn about the strengths, limitations and responsible use of the tools. Learn how to write effective prompts and to identify potential weaknesses in the outputs

Make sure the use of generative AI tools supports user and organizational needs and contributes to improved outcomes for Canadians. Identify appropriate tools for your task; AI tools aren't the best choice in every situation

Potential issues and best practices

- Protection of information
- Bias
- Quality
- Public servant autonomy

- Legal risks
- Distinguishing humans from machines
- Environmental impacts

Each issue includes a description of the potential risks and identifies best practices to manage them.

Digitization can and must become a powerful force for societal good—for fairness, inclusion, economic growth, expanded opportunities, innovation, and above all else, equality.

- Orly Lobel, *The Equality Machine*