Statistics Canada's Framework for Responsible ML

Saeid Molladavoudi, Senior Data Science Advisor

ML 2022

August 2022

Delivering insight through data for a better Canada

Outline

1. Context - Why do we need a framework?
2. Statistics Canada's framework
3. Implementation – Peer-review processes
4. Current and forward looking initiatives
5. Research and Innovation
6. Conclusion

Context

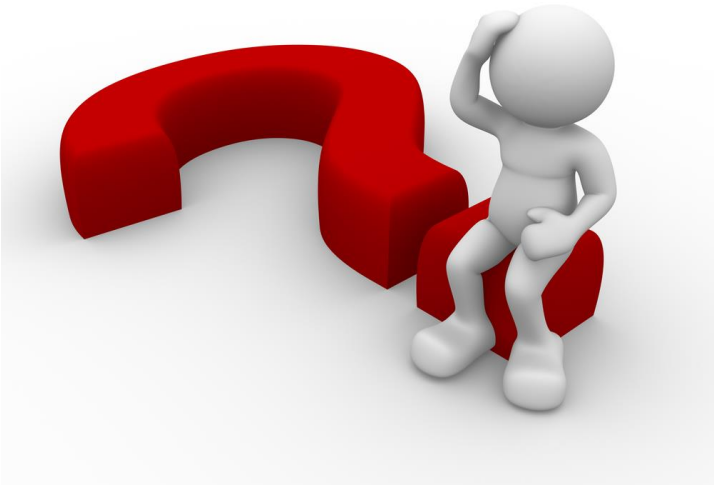
- The use of artificial intelligence and machine learning (ML) methods is continuously increasing inside and outside of Statistics Canada
  - Automate manual tasks (cognitive automation)
  - Develop new products, provide new insights and analysis
  - Use of large amount of data and unstructured data

Machine Learning Word Cloud Stock Photo - Illustration of knowledge, communication: 179040510 (dreamstime.com)

- Canadians expect Statistics Canada
  - To demonstrate **trustworthiness**
  - Produce **quality** statistics (Statistics Canada Quality Framework)
  - Preserve **confidentiality**
  - To use means **proportionate** to the needs (Necessity and proportionality framework)
  - Ensure **transparency**

Context: Why is Responsible ML required?

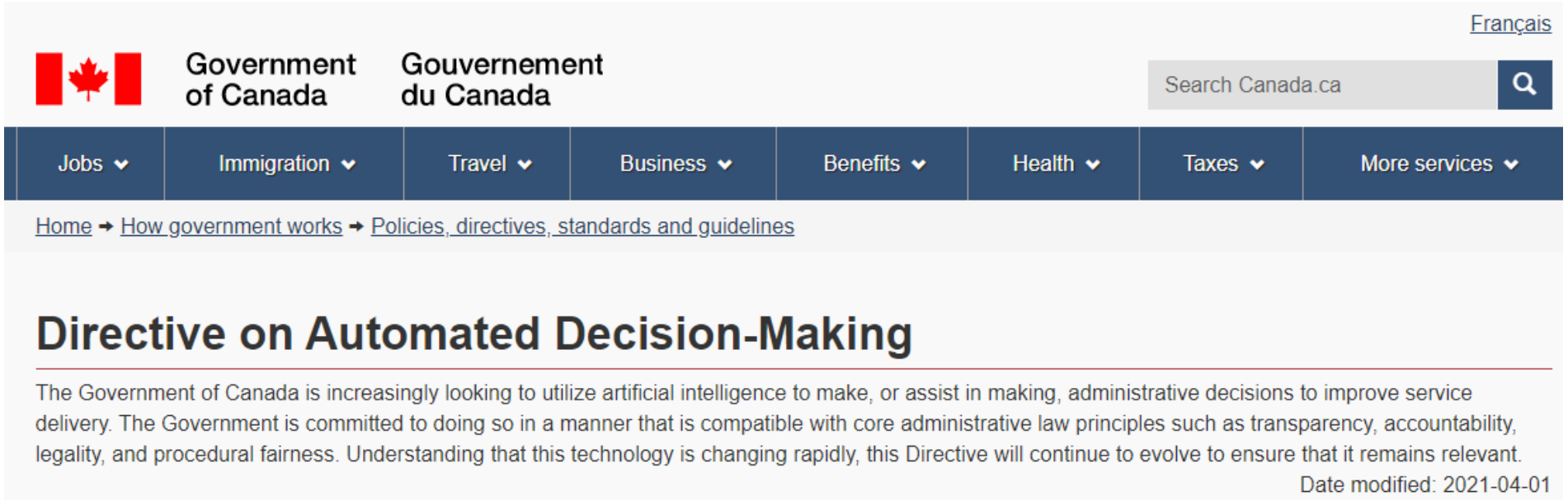
- With the use of these complex methods come many questions
  - What is the **impact** of a decision made by ML methods on humans?
  - Is this data used **appropriate** for our needs?
  - How can we evaluate the **quality** of the results?
- ML models have many **sensitive applications**
  - Diagnostic of a particular disease or health condition
  - Attribution of visas to foreigner workers
- …therefore ML models need to be built in a **responsible** and **transparent** way



Context: Why is Responsible ML required in government?

- **Reproducible** and **transparent** modelling is also essential when ML models are used to make **administrative decisions or generate insights** about individuals and society

- Government of Canada's response:
  - Treasury Board **Directive on Automated Decision-Making** and **Algorithmic Impact Assessment** (AIA)



Context: Why is Responsible ML required in Statistics Canada?

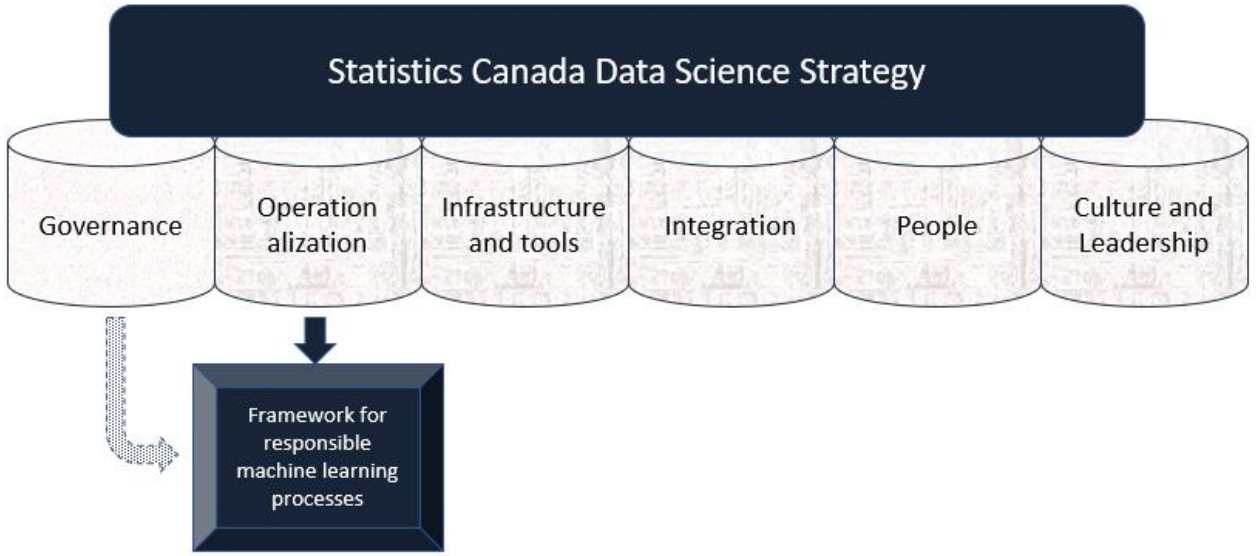
- The responsible ML framework is intimately connected with producing quality statistics and supports the core mission of Statistics Canada:

"The role of Statistics Canada is to **provide credible and relevant statistical information** to the public, to inform decision-making and to help Canadians better understand their country. The trust of the Canadian people is earned and maintained through **transparent and responsible** management of statistical information and of the statistical system that produces it"- **Anil Arora, Chief Statistician, The Quality Assurance Framework, 2017**

https://www150.statcan.gc.ca/n1/pub/12-586-x/2017001/article/s1-eng.htm

Context: Governance for ML at Statistics Canada

- Statistics Canada is about to release its Data Science Strategy based on six pillars
- Operationalization pillar
  - To develop products that are **ethical, safe** and in accordance with Statistics Canada's **confidentiality** and **security** rules.
  - While producing results **faster** and more efficiently, they are continuously subject to **quality, ethics, sound methods and algorithmic accountability.**
- Leading to the use of the Framework for responsible machine learning processes



Framework for Responsible Machine Learning Processes at Statistics Canada

RESPECT FOR PEOPLE

- Value to Canadians
- Prevention of harm
- Fairness
- Accountability

SOUND APPLICATION

- Transparency
- Reproducibility of process and results

RESPECT FOR DATA

- Privacy
- Security
- Confidentiality

SOUND METHODS

- Quality learning data
- Valid inference
- Rigorous modeling
- Explainability

Assessed through self-evaluation and peer review, using a checklist and producing a report or dashboard

Respect for People

Value to Canadians	Prevention of Harms	Fairness	Accountability
• Does the application bring any benefits to users? • How did you show that quality required by the user will be achieved by the ML algorithm used?	Could the results of the machine learning product • Suggest discrimination? • Inadvertently reveal information about vulnerable populations?	• Are all the variables used in the model relevant? • Do you protect integrity and confidentiality of the data? • Is there a strategy in place to avoid personal biases?	• Who is responsible? • Is there human oversight? • What is the plan for monitoring and maintenance of performance of the application?

Respect for Data

Privacy	Security	Confidentiality
Is it compliant with the ***Directive on Conducting Privacy Impact Assessments***?	Is it compliant with the ***Directive on the Security of Sensitive Statistical Information***?	Is it compliant with the ***Policy on Privacy and Confidentiality***?

Sound Application

Transparency

Explain why the algorithm and the learning data are appropriate

- Provide descriptions of the learning data, algorithm, model diagnostics and code
- Share the code if possible

Were all partners, namely subject matter experts, methodologists, data scientists and computer scientists, involved in the development of the model?

Reproducibility of Process and Results

- Is the code version-controlled, e.g. via GitHub or GitLab?
- Are the outputs bundled with the code and data?
- Can the process be executed from a simple master script?
- Is the pipeline fully documented?

Sound Methods

Quality of Learning Data

- Do you have sufficient labelled data?
- How was the quality assessed?
- Is it a good representation of the target population?
- Is there a process to detect data drift?

Valid Inference

- Why is your validation protocol appropriate?
- What are the evaluation metrics?
- Is there a quality assurance methodology in place to track the performance of the model?

Rigorous Modelling

- How does the model perform on never-seen-before data?
- What is the assessment of the generalization error?
  - Overfitting
  - Underfitting

Explainability

- Can you explain the relationship between the input and the output of a model?
- Do you use any tool or methods to aid the interpretation of the model?

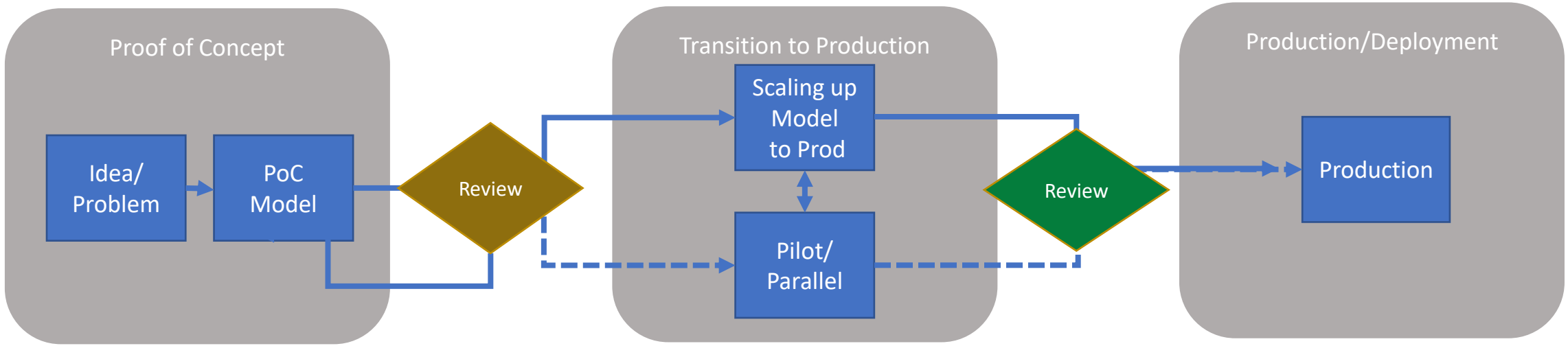
Implementation

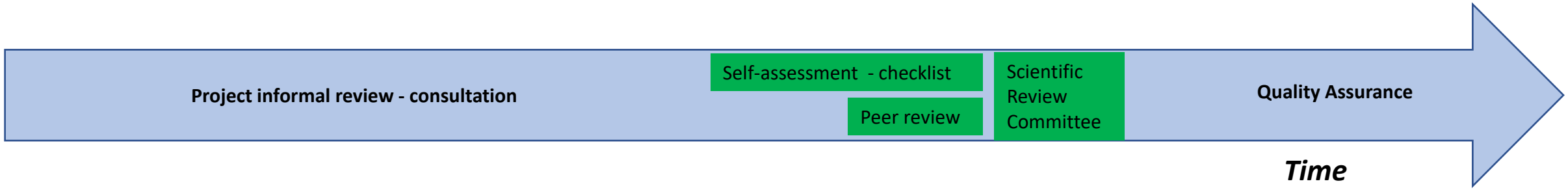
Peer-Review Processes

Implementation – Review Process

- The Framework for responsible ML processes at Statistics Canada is enforced through an **independent** review process of projects for internal and external clients (other federal departments)
- We peer-review proof-of-concept ML projects before they transition to production/deployment.
- Several projects have been evaluated using this review process. Here are a few from last year:
  - Crop Yield Prediction (using satellite data and other)
  - Building Type Classification from Street-view Imagery using Convolutional Neural Networks
  - Census comments classification
  - Receipts auto-Capture (Information Extraction and Optical Character Recognition)
  - (scanned) PDF Information Extraction (financial statements)
  - Client Inquiry Text Classification (classification and automated rerouting of requests/emails)

High level steps from Inception to Production (evolving process)





Implementation – The checklist

The framework for responsible ML processes at StatCan is enforced through an **independent** review process. The process consists of a self-evaluation using a checklist, a peer-review and a presentation to the Modern Statistical Methods and Data Science (MSMDS) branch Scientific review committee.

- **The Checklist**
  - The checklist consists of a list of questions whose honest answers will indicate whether or to what extent the guidelines have been followed.
  - It translates each framework guideline into specific questions
  - To be completed by the project manager/team
  - Could trigger the completion of the Algorithmic Impact Assessment tool
  - Once completed, the list is sent back to the review team

Implementation – The checklist (Cont'd)

- Translates each framework guideline into specific questions.
- Could trigger the completion of the Algorithmic Impact Assessment tool

Sound Application *- Reproducibility of Process and Results*

English Français

Guideline No.	Checklist questions	Yes	No	Write-in	Self-assessment filled by	Peer review conducted by	Peer review comments
18	Is the code version controlled with a corporately supported system (such as GitLab, GitHub)?	○	○				
19	Is the set of modeling results bound with the code, input data and system and session information?	○	○				
20	Is the pipeline fully executable by all stakeholders from a "master" script?	○	○				

Implementation - Review Process & Committees

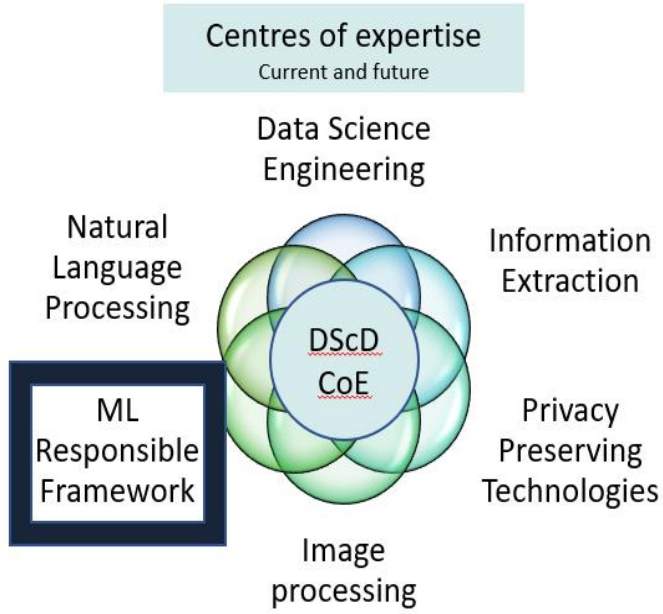
- **Peer review Process**
  - **Designated reviewers** review the completed checklist and relevant documentation provided by the project manager/lead (methodological report, access to code, etc.).
  - **Data Science Division** - Experts in statistics and data science
    - Review themes **Sound Application** and **Sound Methods**
  - **Data Ethics Secretariat** – Experts in data ethics
    - Review themes **Respect for People** and **Respect for Data**
    - Depending on the project, the **Internal or External Ethics Committee** may be called to review the project
  - The appointed reviewers then produce a **report** to the project manager/lead possibly listing major and/or minor recommendations
- **(Possible) Presentation to the MSMDS Scientific Review Committee**
  - Review themes **Sound Application** and **Sound Methods**
  - Provide recommendations and guidance

Recent and forward looking initiatives

- **Raise awareness** within and outside Statistics Canada
  - Presented at the **Advisory Council on Ethics and Modernization of Microdata Access** meeting (June 2021)
  - Article in the **Data Science Network newsletter** (July 2021)
  - Presented at the 2021 **International Methodology Symposium** (October 2021)
    - Quality Considerations in the Production of Statistics
    - Ethics and Privacy workshop
  - Presented at the 2021 **CANSSI Showcase** (Nov. 2021) Workshop: Innovations in Data Analytics and Data Science: What's New at Statistics Canada?
  - Presentation at the event organized by the Data Science Interdisciplinary Research Cluster at the University of Toronto's **Dalla Lana School of Public Health** (April 2022).
- **Develop a workshop/training** to promote responsible AI practices (three-module course)
- **Continue to review** ML processes moving to production using the framework
- **Conduct research** in the areas related to responsible ML, including applications of explainable AI

Recent and forward looking initiatives (Cont'd)

- **Ever-green approach** to the framework and the review process
  - **Publish the checklist**
  - **Review the framework** on a yearly basis to remain relevant
  - Continue to provide tools/support to make it easier to follow the checklist
- **Outreach and collaboration** with other federal departments and leading private sector and not-for-profit organizations
  - Microsoft (best practices for responsible AI)
  - ForHumanity (Independent Audit of AI Systems oversight)
- We're open to other **collaborations**



Research and Innovation

We conduct research in the areas of responsible and ethical AI/ML

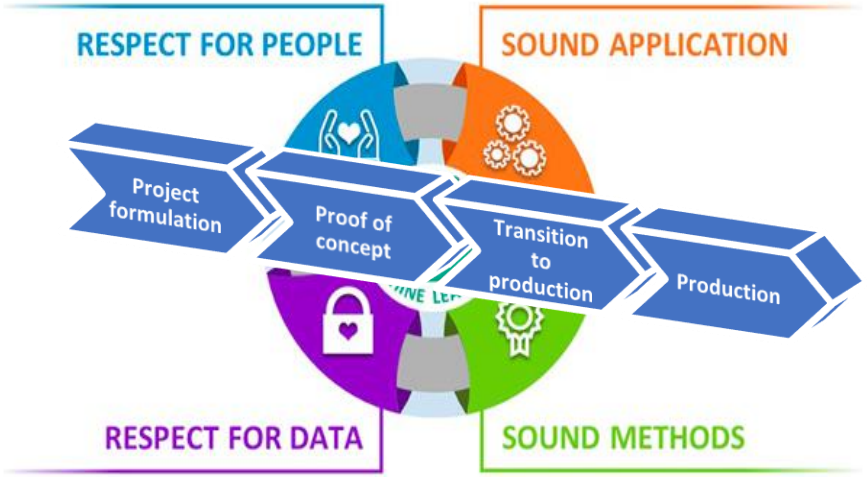
- Research projects on **Explainable AI** (XAI)
  - Study algorithms that help explain AI models
    - Apply methods on relevant projects
    - Identify practical recommendations on method(s) to use
  - Cynthia Rudin's work on interpretability and the use of simpler models for interpretable AI
    - Review of her work in view of practical recommendations for choice of models
- Research project on **Automated Machine Learning** (**AutoML**)
  - Review available tools to create guardrails and recommendations for their use
  - Identify opportunities, weaknesses, limitations, risks, …
- **Upcoming Research Projects** on Fair ML, Adversarial scenarios, Confidentiality and model privacy, Causal ML and counterfactual analysis and others.

Conclusion

- **Needs/Requirements**
  - Use AI and ML methods responsibly
  - Maintaining the trust of Canadians
  - Produce good quality statistics
- **Statistics Canada's Response**
  - Framework and review process

- The framework promotes responsible development by flagging potential design vulnerabilities

- The framework also gives developers and reviewers a benchmark for responsible collaboration



For more information,

please contact my team:

Loic Muhirwa
Loic.muhirwa@statcan.gc.ca

Mohammed Haddou
Mohammed.Haddou@statcan.gc.ca

Étienne Rassart
Etienne.Rassart@statcan.gc.ca

Saeid Molladavoudi
Saeid.Molladavoudi@statcan.gc.ca

The content of this presentation represents the position of the authors and may not necessarily represent that of Statistics Canada.

Pour plus d'information,

veuillez contacter mon equipe:

Loic Muhirwa
Loic.muhirwa@statcan.gc.ca

Mohammed Haddou
Mohammed.Haddou@statcan.gc.ca

Étienne Rassart
Etienne.Rassart@statcan.gc.ca

Saeid Molladavoudi
Saeid.Molladavoudi@statcan.gc.ca

Le contenu de cette présentation représente l'opinion des auteurs, mais pas nécessairement celle de Statistique Canada.