Data Ethics and Responsible AI at DFO

Guidance for ethical collection, use and sharing of data

Why Does Data Ethics Matter?

- With data increasingly seen as a strategic asset, activities related to the collection, usage and sharing of data are integral to almost every project at the Department of Fisheries and Oceans (DFO)

- How we approach these data-centric activities can have a major impact on the outcomes of a project and on other projects that depend on the project deliverables

- Unethical and irresponsible handling of data assets can have a broad range of consequences:
  - Decision-making and policies resulting in harms (e.g., stigma, financial loss, etc.) to communities
  - Violations of personal privacy
  - Loss of public trust

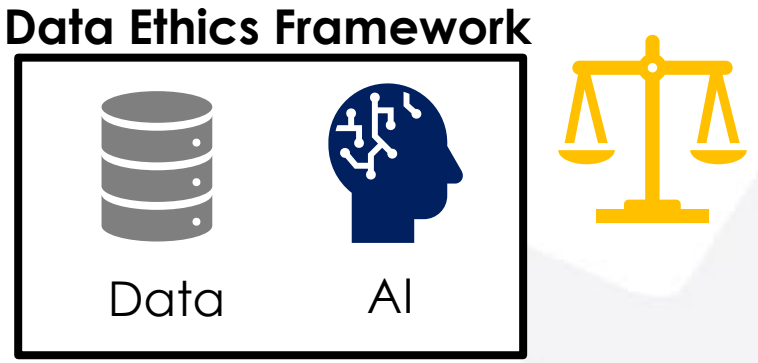
- Conversely, careful handling of data supported by sound ethical guidance can result in more positive project outcomes and greater benefits to the public

- A data ethics framework is a key part of realizing the departmental Data Strategy objective for Data as an Asset:
  - *"DFO values data as a strategic asset that is fit for use, discoverable and shareable, and is used appropriately, effectively and ethically to deliver on the Department's core responsibilities and achieve results for Canadians."*

Building a Data Ethics Framework

- DFO has begun work to develop a departmental data ethics framework to provide guidance on the ethical responsibilities of handling data

- The framework covers two inter-related domains:
  - Ethical handling of data
  - Ethical use of artificial intelligence (AI)



- The framework is structured into the following core sets of components:
  - **High-level guiding principles** defining the goals and standards set out by the framework
  - **Concrete processes** supporting the realization of the guiding principles
  - **Documentation and forms** to complement and facilitate the processes

Establishing Guiding Principles

- Guiding principles cover core themes and provide high-level goals that are relatively agnostic to project-specific details

- Establishing a set of guiding principles helps to orient the framework towards the areas in which efforts should be focused
  - Supporting processes can then be developed to fill in the specifics on how to attain the goals set out by the guiding principles

- The guiding principles for the DFO Data Ethics Framework have been identified through the study of:
  - Guiding principles as defined in existing data ethics frameworks from other sources
  - Ethical questions explored in DFO project use cases

Comparison of Data Ethics Guiding Principles

	Public Benefit	Security and Privacy	Transparency	Accountability	Methodology and Quality	Other
Statistics Canada	• Benefit to Canadians • Report Fairly and Do No Harm	• Secure and Private Use	• Transparent and Accountable Processing	• Transparent and Accountable Processing	• High Quality Data	• Trustworthy and Sustainable Processing
United States	• Respect the Public, Individuals and Communities	• Respect Privacy and Confidentiality	• Promote Transparency	• Uphold Statutes, Regulations and Standards • Hold Oneself and Others Accountable		• Act With Honesty, Integrity and Humility • Stay informed of Developments in Data Management and Science
Germany	• Human Dignity • Democracy • Justice and Solidarity	• Privacy • Security				• Self-determination • Sustainability
United Kingdom	• Fairness		• Transparency	• Accountability		
UK Statistical Authority	• Public Good • Public Views and Engagement	• Confidentiality, Data Security	• Transparency	• Legal Compliance	• Methods and Quality	
Canada's Digital Charter	• Data and Digital for Good	• Safety and Security	• Transparency, Portability and Interoperability	• Strong Enforcement and Real Accountability		• Others…

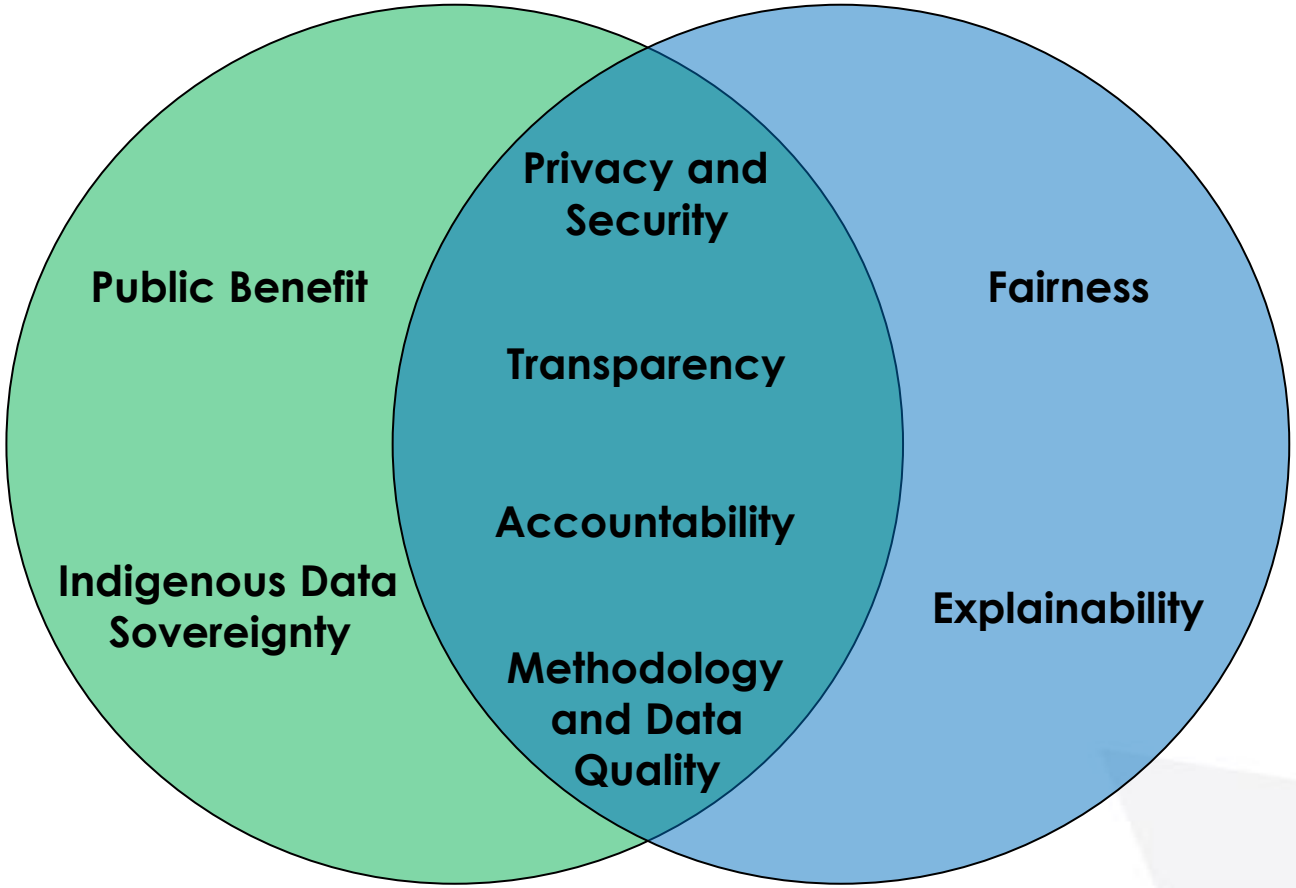
Comparison of Responsible AI/ML Guiding Principles

	Sound Methodology	Transparency	Accountability	Fairness	Privacy and Security	Value and Benefits	Interpretability /Explainability	Other
Government of Canada	• Understand and Measure	• Be Transparent • Be As Open As We Can					• Provide Meaningful Explanations	• Provide Sufficient Training
Statistics Canada	• Reproducibility of Process and Results • Quality Learning Data • Valid Inference • Rigorous Modeling	• Transparency	• Accountability	• Fairness	• Privacy • Security • Confidentiality	• Value to Canadians	• Explainability	• Prevention of Harm
Gartner		• Explainable and Transparent	• Accountable	• Fair	• Secure and Safe	• Human-Centric and Socially Beneficial	• Explainable and Transparent	• Human-Centric and Socially Beneficial
Microsoft	• Reliability and Safety	• Transparency	• Accountability	• Fairness	• Privacy and Security			• Inclusiveness
European Commission	• Technical Robustness and Safety	• Transparency	• Accountability	• Diversity, Non-Discrimination and Fairness	• Privacy and Data Governance	• Societal and Environmental Well-being		• Human Agency and Oversight
US Government Accountability Office	• Ensure Quality Reliability and Representativeness of Data Sources and Processing • Ensure Reliability and Relevance Over Time		• Promote Accountability			• Produce Results that Are Consistent With Program Objectives		

Selected Themes for DFO Guiding Principles

Data Ethics

The principles of data ethics apply to all projects that interact with data.



Responsible AI

Responsible AI is an extension of data ethics. All the principles of data ethics apply, some are extended in scope, and some new principles are added.

Within each of the identified themes, we have established a set of guiding principles as tabulated guidelines

Framework Guidance Materials

- The core of the data ethics framework will be composed of the following guidance materials currently under development

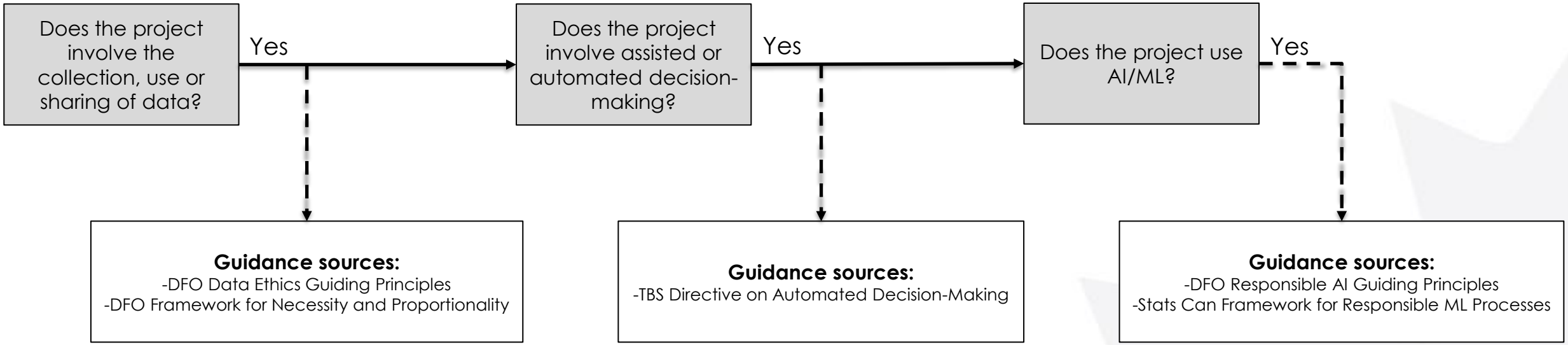
  - **Guiding principles:**
    - Guidelines document

  - **Processes:**
    - Framework for Necessity and Proportionality (adapted from Statistics Canada)
    - Framework for Responsible Machine Learning Processes (adapted from Statistics Canada)

  - **Supporting documents:**
    - Data quality assessment guidelines
    - Framework for Necessity and Proportionality form
    - Framework for Responsible Machine Learning Processes form

Identification of Guidance Sources

Project teams can follow a simple decision tree to identify which guidance materials from the framework pertain to their project



Identification of guidance sources along the path is cumulative

Path Forward to Operationalization

- Operationalization of the data ethics framework requires additional resources to enable effective use of the guidance materials

- The first iteration of the framework will offer a web-based tool for self-assessment
  - Project teams will use the tool to identify guidance material and relevant process steps
  - The tool will allow teams to self-assess their project and receive assessment results and feedback

- Future iterations of the framework will seek to identify expertise in the fields of data ethics and responsible AI to support:
  - Independent review processes
  - Advisory boards
  - Training materials and programs

Office of the Chief Data Steward / Bureau de l'Intendance principal des données

Strategic Policy
Politiques Stratégiques

NCR_OCDO-BDPD_RCN@dfo-mpo.gc.ca