The Role of Data Ethics to Maintain and Improve Public Trust: The Statistics Canada Experience

Martin Beaulieu,
Chief - Quality and Data Ethics Secretariat, Statistics Canada
Workshop on Ethics in Modern Statistical Organisations
March 27, 2024

Delivering insight through data for a better Canada

Statistics Canada / Statistique Canada

Canada

Outline

- What is Data Ethics?

- Why is it important for National Statistical Organizations?
  - Public Trust

- Statistics Canada's Data Ethics Approach
  - Necessity and Proportionality Framework
  - Guiding Principles
  - Concrete Implementation
  - Challenges and Lessons Learned

Statistics Canada Statistique Canada

Canada

Data Ethics

- Field of study that addresses questions about the appropriate use of data, at various stages of the life cycle.

- Examples of possible ethical issues related with the use of data:
  - ➢ Asking questions that could potentially cause harm to respondents
  - ➢ Using biased data to inform decision makers
  - ➢ Using data that can reinforce existing inequalities
  - ➢ Privacy intrusion that is not proportional to the benefit of a project
  - ➢ Producing information that does not correspond to the needs of the population

Statistics Canada Statistique Canada

Canadä

Importance of Data Ethics

- Ever increasing demand for more timely and more granular statistics
  - May lead to surveys asking more sensitive questions
- New data sources available and new tools to process these data combined with a constant decline in surveys response rates
  - In this context, with we are innovating and therefore introducing practices that can raise new ethical challenges.
  - Less transparent by default
- Impact on Public Trust
  - Data Users
    - Quality
  - Data Providers (population, enterprises, other organizations)
    - Confidentiality/Security
    - Responsible acquisition and use throughout the lifecycle
      - Gather – Guard – Grow – Give (Rancourt, 2019)

Necessity and Proportionality Framework (N&P)

- Objective to optimize privacy protection and the production of information when designing a data-gathering approach.
- **Necessity**: The description of the needs should include **who** requires the information, the reasons **why** such information is needed and concrete expected benefits.
- **Proportionality**:
  - The balance between privacy intrusion and the quality of the data is coherent with the expected benefits of a project;
  - Ensuring we do not acquire or use more information than we need to produce such benefits.
- Concrete application of the N&P Framework:
  - Creation of the Data Ethics Secretariat
  - Foundations of ethical considerations: Six Guiding Principles

Foundations of Ethical Considerations:
Six Guiding Principles

Benefits for Canadians

Privacy & Security

Transparency & Accountability

Trust & Sustainability

Data Quality

Fairness & Do no Harm

Guiding Principles

Benefits for Canadians

- Define and illustrate the expected benefits with concrete examples
- If not going through with the project has negative side-effects, explaining them and listing the consequences will help justify the project.

Privacy & Security

- Measures to avoid a security breach/Potential impacts of a breach
- Gathering only the necessary information

Transparency & Accountability

- Proactive steps taken to be transparent
- Accountability when using ML or AI
  - ➤ "Human in the loop"

Guiding Principles

Trust & Sustainability

- Relationships with partners/data providers
- Ensure our own rules are followed

Data Quality

- Known sources of error and/or bias and ways to address them
- The expected benefits of the project should be proportional to the quality of the data.

Fairness & Do no Harm

- Making sure harmful profiles that could propagate stereotypes are not created
  - Especially important with information on vulnerable or marginalized segments of the population
- Data to be used for the purpose of the project that justified its acquisition
  - Plans to reuse data for different purposes

Examples

- Personal questions responded by proxy (Privacy and security)
- Difficult survey questions (fairness and do not harm; data quality)
- Using administrative data to compensate for refusals (Transparency and Accountability; Trust and sustainability)
- Linking multiple data sources to create very precise profiles

Concrete implementation

- Reviews are done by Data Ethics Secretariat (DES) on new data acquisitions, surveys and microdata linkages, working alongside other groups within the agency:
  - Legal/Privacy team
  - Subject-Matter area
  - Specialized teams depending on context
- Short questions based on the six guiding principles used by Data Ethics Secretariat to determine whether a request/project is sensitive.
- Integrated existing processes

Governance Structure

- Data Ethics Secretariat

- Internal Data Ethics Committee

- Principal Data Ethics and Scientific Integrity Officer

- [Advisory Council on Ethics and Modernization of Microdata Access](#)

Challenges and Lessons Learned

Challenges

- DES role not always well known
- Some definitions not well known (e.g. personal data)
- Perception of adding a roadblock
- Time management for a small team

Lessons Learned

- ***COMMUNICATION***
- Support from governance structure
- Get involved early
- Feasibility/Exploratory studies
- Don't get rushed by tight timelines

Thank you/ Merci

Pour de plus amples renseignements, veuillez contacter:

For more information, please contact:

martin-j.beaulieu@statcan.gc.ca

Statistics Canada / Statistique Canada

Canada