Adapting Quality Assurance Frameworks to the Fast-Evolving World of Official Statistics

Martin Beaulieu and Roxanne Gagnon

Quality and Data Ethics Secretariat, Statistics Canada

11th European Conference on Quality in Official Statistics (Q2024)

June 5, 2024

Delivering insight through data for a better Canada

Introduction

- Official statistics are essential for organizations (government, private sector, civil society, etc.) to make data driven evidence-based decisions.
- Statistical ecosystem has changed significantly:
  - More data sources
  - New technologies available
  - More data producers occupying statistics space
  - More data/information needs
    - Need for more timeliness
    - Need for more granularity
  - More survey related challenges
    - Declining response rates
    - Response burden
    - Cost

Introduction

- National Statistical Organizations (NSOs) have earned the reputation of producing high quality statistics.
- Most NSOs have developed Quality Assurance Frameworks (QAFs), to define and describe the measures taken to manage quality.
  - Statistics Canada developed its first QAF in 1997. The latest version was released in 2017
- Most of these QAFs were developed in a context where NSOs had full control of the entire statistical process
- With changes in the recent years, NSOs are no longer only data producers, they are also data users
  - Data could be ingested or shared at any point during the data lifecycle
- While their content remains valid, QAFs may need to be expanded to account for the new context.

QAF revision – 3 key elements

- NSOs have become data users
  - Data Stewardship role
  - Quality assessment of data coming from alternative sources
- The importance of social acceptability and the trust in NSOs
  - Ethical use of data (data ethics)
  - Necessity and Proportionality
  - Clear principles to demonstrate trustworthiness
- Modelling using ML algorithms
  - Balance between accuracy and explainability/reliability/inferential quality

QAF revision – Data Stewardship

- Existing QAFs cover well the data stewardship activities within the NSOs.
- With NSOs becoming data users, the importance of good data stewardship outside of the NSO has increased
- The promotion of the use of standards is essential to facilitate data sharing and reuse.
  - NSOs have a role to play in the promotion of standards.
  - To what extent can Statistics Canada, as an NSO, influence other organizations within the public or even the private sector?
- Data quality for non-statistical sources: Government of Canada Data Quality Framework

QAF revision – Social acceptability/Trust

- As the number of potential data sources increases, social acceptability/trust is essential for NSOs to benefit from them
- Data holders will be more open to sharing data with NSOs if they know that individuals accept this
- Frameworks on Data Ethics, Necessity and Proportionality and confidentiality have been developed to gain this acceptance/trust
- Application of data ethics principles is not new
  - Increase in use of alternative data sources leading NSOs to be more transparent about data acquired and their use.

QAF revision – Social acceptability/Trust

- Necessity and Proportionality Framework
  - **Necessity**: The description of the needs should include **who** requires the information and the reasons **why** such information is needed.
  - **Proportionality**: The balance between privacy intrusion and the quality of the data is coherent with the expected benefits of a project; and that we do not use more information than we need to in order to produce such benefits.
- In practice, reviews of new acquisitions or uses are founded on six guiding principles

Statistics Canada Data Ethics Six Guiding Principles:

Benefits for Canadians

Privacy & Security

Transparency & Accountability

Trust & Sustainability

Data Quality

Fairness & Do no Harm

QAF revision – Modelling

- Model explainability is not a new issue, but there is renewed interest with the advent of AI/ML and increasing needs for modelling
- How should we approach issues related to complex models that may be difficult to explain?
- Finding the right balance between accuracy and explainability, reliability and inferential quality is a challenge for NSOs.
- Some considerations:
  - Solid model validation processes are crucial
  - Clear understanding of users and their needs is key

    →When relations between factors need to be understood, simpler models may be required

    →When predictions are needed, precision (and complex models) may be required
  - Explainability is a form of the interpretability quality dimension

QAF revision – Key elements vs current principles and quality dimensions

Key Element	QAF Principles	Quality Dimensions
Data Stewardship	Managing relationships with data users, data providers and other stakeholders Managing statistical standards	Coherence and comparability Interpretability and managing metadata
Social Acceptability/ Data Ethics	Transparency Commitment to quality	Accuracy and reliability Relevance
Model Explainability	Methodological soundness	Accuracy and reliability Interpretability and managing metadata

Statistics Canada Quality Dimensions
Relevance
Timeliness
Accuracy
Accessibility
Interpretability
Coherence

Summary

- Statistics Canada's upcoming new QAF should be a foundation to assure quality remains a core value
- A well-established culture of quality management, data stewardship and data ethics will help maintain and improve the trust of Canadians towards their NSO and official statistics

Thank you/ Merci

Pour de plus amples renseignements, veuillez contacter:

For more information, please contact:

martin-j.beaulieu@statcan.gc.ca
roxanne.gagnon@statcan.gc.ca

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