

GC Data Quality Framework

The GC Data Quality Framework has been adapted and formalized into the [Guidance on Data Quality](https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/information-management/guidance-data-quality.html) (<https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/information-management/guidance-data-quality.html>). This page is no longer supported by Treasury Board of Canada Secretariat but remains available online for reference purposes. This page may contain useful information or guidance, however some links or references may be out of date.

Contents

Background

Purpose

Overview

Framework

Guidelines

Appendix A: Glossary of Terms

Appendix B: Examples of Applications

Appendix C: Approach

Appendix D: References

Background

Data is foundational to digital government. The Government of Canada (GC) increasingly relies on data to support the design and delivery of programs and services. Data is also a critical building block of evidence-informed policymaking, enabling the government to make measured and timely decisions that benefit all Canadians. It also supports the government's commitment to openness and transparency, helping build public trust in digital government. Data also plays a role in advancing international cooperation and helping Canada meet its international obligations.

For data to be effective and trustworthy, it needs to be fit-for-purpose. Fitness-for-purpose is an indicator of data being both usable and relevant to user needs and goals.^[1] The quality of data has a significant impact on its value to users. It influences whether data is discoverable and available to users when they need it, and the ways in which they can use or reuse data within and across organizations and jurisdictions. The prominent role of data in government operations and decision-making also highlights the importance of high-quality data not only to the government's mandate, but also to public trust. Inaccurate or incomplete data, for example, can lead to misguided policies or biased decisions with adverse impacts on individuals, communities, or businesses. Managing the quality of data throughout the lifecycle – from acquisition to disposition or archiving – can help ensure it is fit-for-purpose, allowing users to appropriately harness its value to support their objectives. This draws on multiple roles in an organization: for example, data providers and custodians ensure data is managed to be usable, while data stewards and consumers determine its relevance within a specific use-context.

There is a need for a common understanding of data quality in the federal government. The current landscape includes a wide range of approaches to data quality, each developed to suit a specific type of data or organizational context. While such focused approaches serve a unique function, a shared framework with broad applicability can strengthen government-wide data governance capabilities by establishing a common vocabulary, improving coherence in data quality rules, facilitating interdepartmental data sharing and reuse, and fostering trusted data flows and ethical practices.

The GC Data Quality Framework (the Framework) is a response to the growing need for central direction in this area. Data quality has recently emerged as a government-wide priority. The Treasury Board (TB) Policy on Service and Digital holds the GC Chief Information Officer (CIO) responsible for prescribing an enterprise-wide standard on data quality (subsection 4.3.1.1). The 2021-2024 Digital Operations Strategic Plan identifies the need for a government data quality framework as a priority action. Similarly, Recommendation 17 in the Data Strategy Roadmap for the Federal Public Service (Data Strategy Roadmap) calls for the creation of an adaptable government-wide data quality framework. The need for common direction on data quality is also evident in Budget 2021, which includes various investments in data capabilities across priority areas such as health, quality of life, justice, and the environment. This is also reflected in recent ministerial mandate letters, which commit cabinet ministers to improving the quality and availability of disaggregated data to foster fair and equitable policymaking.

Data quality is also being prioritized within federal departments and agencies. Many departmental data strategies developed following the publication of the Data Strategy Roadmap identify data quality as an organizational priority and list planned or existing efforts aimed at managing it effectively. Further, the TB Directive on Service and Digital requires departmental CIOs and other designated officials to ensure that “information and data are managed to enable data interoperability, reuse and sharing to the greatest extent possible within and with other departments across the government to avoid duplication and maximize utility, while respecting security and privacy requirements” (subsection 4.3.1.3).

The concern with data quality also extends to automated decision systems, which rely on data to perform their functions. The TB Directive on Automated Decision-Making requires federal organizations to validate the quality of data collected for and used by automated decision systems (subsections 6.3.1, 6.3.3). The Algorithmic Impact Assessment, a risk assessment tool that supports the Directive by determining the impact level of an automated decision system, also accounts for this by asking users to identify processes for testing bias in data. Taken together, these measures are part of a broader move towards treating information and data as strategic assets “to support government operations, service delivery, analysis and decision-making” (Policy subsection 4.3.2.1).

The COVID-19 pandemic has amplified the need for government-wide approaches to data governance in the GC. Tracking, analyzing, and controlling the spread of the virus in Canada has required the government to mobilize data collection, sharing, integration, and reuse capabilities in collaboration with provincial, territorial, and international partners. The effectiveness of this operation depends on the ability of GC organizations to acquire accurate and timely data for (dis)aggregation and analysis. A shared understanding of such quality concepts across government can enhance horizontal data capabilities, bolster the GC’s pandemic response, and improve public trust.

Purpose

The purpose of the Framework is to establish a government-wide approach to the definition and assessment of data quality. This will support whole-of-government priorities, digital policy goals and requirements, and user needs by:

- Supporting compliance with the TB Policy and Directive on Service and Digital by informing enterprise-wide and departmental approaches to data and information quality;
- Enabling consistent approaches to the assessment of enterprise data and information quality, including in the context of open data, enterprise architecture, and government-wide data and information governance; and
- Supporting strategic data priorities identified in the Data Strategy Roadmap, Digital Operations Strategic Plan, and ministerial mandate letter commitments.

The Framework aims to strengthen government-wide capabilities in data quality management and control with a view to:

- Improving the availability, interoperability, usability, and public value of data;
- Facilitating data sharing and reuse;
- Supporting the use of data analytics; and
- Building trust in data

These objectives will help advance evidence-informed decision making and enhance the design and delivery of policies, programs, and services across government.

Overview

The Framework defines data quality in terms of nine dimensions: Access, Accuracy, Coherence, Completeness, Consistency, Interpretability, Relevance, Reliability, and Timeliness. Data can be considered fit-for-purpose to the degree that it satisfies these criteria. The Framework is intended to apply to all data types and use (or reuse) contexts. It is also technology-agnostic.

The dimensions provide users with a conceptual vocabulary for identifying and analyzing a broad range of intrinsic and extrinsic data quality issues to ensure that data is usable and relevant to user objectives. Common issues include challenges comparing datasets obtained from multiple sources, delays in acquiring time-sensitive data, and inaccuracies in client information. (See Appendix B for examples specific to each Framework dimension.)

The dimensions are not mutually exclusive; they overlap in practice as data quality issues tend to be multifaceted – a dataset with multiple representations for the same concept, for example, could be both incoherent (difficult to integrate or compare with other datasets) and uninterpretable (difficult to understand). However, the emphasis placed on each dimension could vary based on user needs, which could necessitate trade-offs. Time-sensitive data needs, for example, could lead users to accept compromises in accuracy in order to ensure timeliness. It is also important to recognize that not all dimensions will necessarily be applicable to a use-case. Prioritization could also depend on the lifecycle stage under consideration.

Each Framework dimension is supplemented with guidelines that enable users to interpret and apply it. The guidelines identify actions that can inform and standardize approaches to data quality assessment. While they may not necessarily be relevant all at once, they can serve as an adaptable checklist for identifying relevant policy and legal requirements, resourcing

considerations, best practices, and stakeholders. Collaboration between organizations can support the implementation of the guidelines, particularly for data shared, reused, or released to the public.

The applicability of the Framework to all types of data provides users with various opportunities for adopting or adapting it to suit their needs. For specific domains of data, the Framework could serve as a base for extensions that either add to the dimensions (horizontal extension) or further elaborate them to support their application in such contexts (vertical extension).

The following is an illustrative list of instruments and governance processes that could benefit from the common direction on data quality established in this Framework:

- Management Accountability Framework (MAF) (e.g., in assessments of departmental maturity in lifecycle data management);
- GC Digital Standards (e.g., in assessments of digital initiatives against the “be good data stewards” standard);
- Algorithmic Impact Assessment (e.g., as supplemental guidance to questions pertaining to data quality frameworks and processes);
- Privacy Impact Assessment (e.g., as supplemental guidance to assess the privacy impacts of programs or activities involving personal information, which includes considerations pertaining to its accuracy);
- GC Enterprise Architecture Framework (e.g., in assessments of digital initiatives against the information architecture layer of this framework);
- Departmental data policies and related data quality frameworks and tools (e.g., in requirements, principles, governance structures, or business rules related to data quality);
- Interdepartmental or intergovernmental data sharing agreements (e.g., in clauses establishing quality provisions for data being shared or exchanged); and
- TB submissions (e.g., as a common vocabulary for articulating data quality issues and objectives in the context of a program’s design or implementation).

Framework

Data can be considered fit-for-purpose to the degree it satisfies the following dimensions. The dimensions are principles describing intrinsic and extrinsic aspects of data quality in the government.

Access: The ease with which data can be discovered, processed, manipulated, and obtained by a user.

Access is a measure of how available and ready data is to meet user needs. This depends on several factors such as whether users are aware of the data and able to gain authorized access to it. Even when accessed or acquired, however, users may not always have the capacity to process or manipulate it to meet their needs due to technical, resource, informational, policy, or legal limitations.

Accuracy: The degree to which data describes the real-world phenomena it is intended to represent.

Data is accurate when it represents a phenomenon adequately. Assessments of accuracy vary by context, methodology, and the validity of underlying hypotheses or assumptions. Maintaining accuracy in public sector organizations involves ensuring that data collected to administer

services matches what clients shared. In policy and program initiatives, ensuring accuracy often requires users to validate data by consulting trusted sources and evaluating the methods or processes by which data was acquired.

Coherence: The degree to which data from one or more sources is comparable and linkable.

A coherent dataset conforms to common architecture taxonomies or classifications. Users can improve data's coherence by adopting applicable organizational, federal, national, or international standards. Coherent data is reusable and interoperable; users can also integrate and compare it with other data.

Completeness: The degree to which data values are sufficiently populated.

Data can be considered complete when it has the entries needed for users to use it appropriately. Contextual and substantive information enables users to make sense of a dataset in their respective lines of business.

Consistency: The degree to which data is internally non-contradictory.

Consistency helps ensure the logical validity of a dataset. A dataset is consistent if the relationships linking its components are determined to be logically sound.

Interpretability: The degree to which data can be understood in its appropriate context.

A dataset is interpretable if a user is able to understand its entries, determine why and how it was collected or created, and judge its relevance to a policy, program, service, or other government initiative.

Relevance: The degree to which data is deemed suitable to support an objective.

The relevance of data depends on whether it provides informational or analytical value to support a user objective. Assessments of relevance are context-dependent: The same data could be relevant in one use-context and irrelevant in another.

Reliability: The degree to which variability in data can be explained.

Reliability is about data meeting user expectations over time. A dataset is reliable if users can explain how it evolves or changes over time.

Timeliness: The amount of time between the end of the period to which data pertains, and the time at which that data is available to meet user needs.

Timeliness is a measure of the delay between two time points: the time when data has passed its reference period and the time when that data becomes available to users.

Guidelines

The guidelines enable users to interpret and apply the nine dimensions consistently. They identify actions that can inform approaches to data quality assessment. Users are encouraged to identify contact points (e.g., data steward, data custodian, data provider, subject matter expert) who have the appropriate expertise to address inquiries related to each dimension.

Access

- Develop an inventory or catalogue of datasets used to support policy, programs or services.
- Develop metadata describing concepts, variables, and classifications in your data assets in accordance with the Treasury Board (TB) Standard on Metadata and Standard on Geospatial Data.
- Establish processes for documenting, retaining, publishing, archiving, and disposing of data collected or created in your organization.
- Assign security categories to data assets as required under the TB Directive on Security Management.
- Define access rights and privileges for data assets to guard against unauthorized access in compliance with the TB Directive on Security Management.
- Ensure processes and procedures exist to support the production of data in response to requests for information under the Access to Information Act and Privacy Act.
- Ensure that the institution has parliamentary authority to collect or create the data for an operating program or activity, as per the TB Directive on Privacy Practices.
- Use plain language (e.g., as described in the Canada.ca Content Style Guide) and machine-readable formats (e.g., CSV, XML, JSON) to improve data portability and facilitate user processing, manipulation, consumption, publication, and archival.
- Invest in data infrastructures to provide easy and secure access to data in accordance with the 'cloud-first' approach established in the TB Directive on Service and Digital. Sensitive data (Protected B, Protected C, or Classified) should be held in systems located within the geographic boundaries of Canada or within GC organizations abroad (see the Direction on the Secure Use of Commercial Cloud Services and GC Security Control Profile for Cloud-based GC Services for guidance on the secure use of cloud services).
- Provide multiple data access and extraction methods to users. This could include making data available in multiple formats and through accessible APIs developed in accordance with the Government of Canada (GC) Standards on APIs.
- Work in the open by default and publish data to the Open Government Portal in accordance with the TB Directive on Open Government and as permitted within applicable federal privacy, security, and intellectual property frameworks. Using plain language, populate the open data registration record with the required metadata when publishing data.
- Conduct surveys to identify barriers to data discovery, access, and use within your organization.
- Report any unauthorized access or use of data to designated security officers and, where personal information is involved, to the Treasury Board of Canada Secretariat and the Office of the Privacy Commissioner of Canada as required under the TB Directive on Privacy Practices.

Accuracy

- Consult with trusted data sources to identify sources of error, verify content, and understand the context surrounding the data.
- Ensure that data includes standardized metadata to enable users to evaluate data accuracy. Relevant metadata could include information about the source, purpose and method of collection, processing, revisions, coverage, and data model and related assumptions.
- Ensure that data is adequately representative of any domains (e.g., geographic areas, populations) contained within it, as appropriate.
- Adhere to expected value ranges to maintain validity. Explanations for outliers should be provided to data users.
- Develop business rules to validate data for errors consistently, including duplication within a dataset. Apply applicable business rules throughout the lifecycle of data, particularly during data collection and sharing.
- Ensure that your data production methodology includes steps to minimize biases and statistical errors (e.g., sampling error). (Refer to the Total Survey Error framework for sources

of statistical error and related quality indicators. On bias, see the GBA+ process to inform assessments of systemic inequalities which could manifest in data.)

- Ensure that an authoritative source exists for data, where possible.
- Ensure that the institution has legislated authority for any data collection concerning an identifiable individual and that such collection is directly related to an operating program or activity within the institution. Mechanisms should exist to correct personal information if requested (see the TB Directive on Privacy Practices).
- Validate constructs and related assumptions in consultation with subject matter experts to evaluate the precision of data, or the extent to which it corresponds to what the user intends to capture.

Coherence

- Identify applicable organizational, federal, national, and/or international data standards and document differences in practices. This can be captured as part of a government-wide or departmental standards repository.
- Adopt or adapt applicable data standards, particularly when sharing data with other organizations or publishing data to the Open Government Portal. Key aspects of data standardization include classifications, metadata, formatting, accessibility, syntax, semantic encoding, and language. Relevant standards could be domain-specific, designed for specific types of data (e.g., statistical, geospatial).
- Record selected standards in a data inventory or catalogue, as metadata, or in data sharing agreements. If new standards are developed, document reasons for not using existing and applicable data standards.
- Ensure that data elements are defined, classified, and represented in alignment with common data architectures, in accordance with the GC Enterprise Architecture Framework.
- Ensure that concepts, definitions, and classifications are compatible within and across datasets to allow for data comparison and integration. In addition to the internal data environment, efforts in this area can extend to organizations across the GC and external organizations across sectors and jurisdictions.
- Use concordance tables to show discrepancies and transitions between standards used across data sources.
- Reduce data duplication across datasets to support data integrity.

Completeness

- Ensure that no entries, columns, or rows that are central to the purpose of a dataset are missing or incomplete.
- Keep values, concepts, definitions, classifications, and methodologies up-to-date.
- Assign mandatory and optional labels to columns or rows in a dataset in order to facilitate assessments of completeness.
- Supplement data with the appropriate metadata elaborating the context and purpose of its acquisition. Metadata could also flag privacy, confidentiality, or accuracy considerations impacting completeness.

Consistency

- Develop validation rules for all logical relationships encoded in a dataset. This could include rules formalizing the relationship between two interrelated variables such as age and marital status (e.g., minimum marriageable age constrains permissible marital status categories for individuals below a certain age) or municipality and province (e.g., a municipality must occur within a province).
- Validate the consistency of datasets on a regular basis using the relevant validation rules. Validation processes should be standardized and automated to support efficiency.
- Maintain a record of consistency issues identified through data validation procedures and periodically review validation rules to ensure their adequacy and effectiveness.

- Acquire the appropriate metadata from the data provider to learn about the entity classes of a dataset, the values they are intended to permit, and the relations that hold among them.

Interpretability

- Adopt, adapt, or develop controlled vocabularies to ensure that key concepts are named and defined consistently in a dataset. Alignment with government-wide vocabularies such as the GC Core Subject Thesaurus is recommended.
- Conform to organizational, federal, national, and/or international data standards governing permissible values for elements in a dataset (e.g., reference data, master data). This could include domain-specific standards.
- Develop definitional and procedural metadata, complying with applicable TB policy such as the TB Standard on Metadata and considering the needs of target audiences. Metadata could clarify the purpose of data acquisition and provide information on methodology and security categorization.
- Document information required to meaningfully interpret the data and maintain a clear link between this documentation and the data throughout its lifecycle.
- Ensure that users are informed of the appropriate uses of the data and aware of its limitations.

Relevance

- Establish processes to consult stakeholders on their data needs. This could involve leveraging data inventories or catalogues to identify existing holdings and minimize redundant data collection (see the TB Guideline on Service and Digital for guidance on information and data collection).
- Identify data requirements and sources based on business objectives and user needs.
- Assess and document how data assets meet data requirements in order to gauge their relevance. This could involve tracking the ways in which data assets are used and re-used to advance organizational or government-wide priorities.
- Leverage the results of relevance assessments to inform future data acquisition and related lifecycle management and governance activities.
- Establish criteria to ensure that data acquisition efforts strike an appropriate balance between business needs and privacy and security risks (see Statistics Canada's Necessity and Proportionality Framework). In the case of personal information, data acquisition should be directly related to an operating program or activity in the institution.
- Ensure that data with historical or archival value is appropriately preserved to facilitate indefinite retention and discoverability in order to enable reuse in accordance with the Library and Archives Canada (LAC) Act and supporting policy instruments.

Reliability

- Identify and document sources that can directly or indirectly change a dataset. Sources of change could include the phenomena represented, data collection methods, data capture and storage technologies, data processing platforms, legislative or regulatory measures, policy requirements, and cyber-attacks.
- Ensure that data acquisition and analysis methods are clearly articulated to facilitate third-party validation and maintain the integrity of the data production process.
- Test data collection or creation instruments prior to deploying them, documenting calibrations and accounting for variance in results.
- Maintain a record of changes to your data assets to ensure that users can determine their provenance and trace how they have evolved since their inception.
- Identify and document dependencies among data assets linked within a data architecture or in the context of data analysis.
- Support the compatibility of concepts, definitions, and classifications over time. Specify and explain discrepancies in the way these elements are maintained over time.

- Protect data assets from fraudulent or unauthorized activities that could undermine their credibility and impact confidence in the data provider. This includes defining, implementing, and maintaining security controls to meet IT security requirements, in accordance with the TB Directive on Security Management and TB Directive on Privacy Practices.
- Employ digital preservation approaches to monitor and guard against the deterioration of data assets over the course of their lifecycle. This includes conducting regular data integrity checks (e.g., through the use of hashing or checksums) and documenting any evidence of deterioration in accordance with the LAC Act and supporting policy instruments.
- Report tampering or unauthorized destruction of data assets to designated security officers.

Timeliness

- Identify users' current and anticipated data needs, including considerations of time (e.g., reference periods, legislative or policy requirements, service standards).
- Consult with data providers to assess whether data needs can be met without delay and inform data users of any expected punctuality issues. This could include confirming the data provider's ability to meet timelines established in data sharing agreements.
- Ensure that data providers have a data release schedule that documents the stages of the data production process and accounts for discrepancies and delays (e.g., through contingency planning).
- Publish preliminary data to the Open Government Portal when useful to users, in accordance with the TB Directive on Open Government.

Appendix A: Glossary of Terms

Controlled vocabularies: A list of standardized terminology, words, or phrases, used for indexing or content analysis and information retrieval, usually in a defined information domain.

Data: Set of values of subjects with respect to qualitative or quantitative variables representing facts, statistics, or items of information in a formalized manner suitable for communication, reinterpretation, or processing.

Data quality: A characteristic of data determined based on its Access, Accuracy, Coherence, Completeness, Consistency, Interpretability, Relevance, Reliability, and Timeliness. High data quality is an indicator of fitness-for-purpose, which means that data is both usable and relevant in a primary or other use-context.

Data standards: A set of documented rules or guidelines that enable consistent and repeatable description, representation, structuring, and sharing of data.

Information: Knowledge captured in any format, such as facts, events, things, processes, or ideas, that can be structured or unstructured, including concepts that within a certain context have particular meaning. Information includes data.

Information lifecycle: The planning, collection, creation, receipt, capture, organization, use, re-use, dissemination, maintenance, protection and preservation, disposition, and evaluation of information.

Interoperability: The ability of different types of electronic devices, networks, operating systems, and applications to work together effectively, without prior communication, to exchange information in a useful and meaningful manner.

Metadata: The definition and description of the structure and meaning of information resources, and the context and systems in which they exist. **Personal information:** Information about an identifiable individual that is recorded in any form.

Appendix B: Examples of Applications

The following use-cases are intended to clarify what Framework dimensions mean in practice by providing concrete examples of relevant quality issues, suggesting approaches to evaluate or address the issues, and distinguishing between Framework dimensions.

Dimension	Example of Application
Access	A program developing an automated decision system publishes information about the system in machine and human readable formats to the Open Government Portal. As open information, it is easy to discover and obtain by stakeholders across sectors.
Accuracy	A data custodian updates data on the country of citizenship of a recently naturalized citizen to ensure that it matches their new status in Canada.
Coherence	A provincial address register is standardized so that the province of Ontario is represented as 'ON' in order to enable data interoperability and facilitate data sharing among organizations that have adopted the same standard.
Completeness	A survey administrator follows up with survey respondents requesting the completion of mandatory fields in a satisfaction survey in order to be able to generate a complete dataset.
Consistency	A program delivering an external service identifies and corrects an error in a client's date of birth, which had been set later than their application's date of submission – contrary to established validation rules.
Interpretability	The International Merchandise Trade Database provides clear definitions of key concepts and accessible descriptions of classifications, enabling users to understand and make use of the data in analyses of trends in international trade.
Relevance	A program responsible for retirement pensions collects banking data from applicants after determining the data's role in supporting the processing of benefits payments.
Reliability	Canadian climate data is adjusted to account for shifts due to changes in instruments and observing procedures. For example, rainfall gauge data extracted from the National Climate Data Archive has been adjusted to correct for factors such as wind undercatch, evaporation, and gauge-specific wetting losses.
Timeliness	Provinces and territories report COVID-19 case data to the federal government every 24 hours to support the daily COVID-19 epidemiology update, which provides a summary of COVID-19 cases across Canada and over time.

Appendix C: Approach

The Framework was collaboratively developed by an interdepartmental working group co-led by Statistics Canada and the Treasury Board of Canada Secretariat (TBS). The group was established in Fall of 2019 under the GC Enterprise Data Community of Practice. The development of the dimensions was informed by an environmental scan of data quality frameworks in the federal government, industry, international organizations, and public sector organizations in other governments.

Regular deliberation among working group members also helped refine the approach to framing and defining the dimensions, while helping build consensus around the framework as a whole. Once there was broad agreement on the foundations of the framework, the group collaborated on

the development of guidelines supporting the consistent interpretation and application of the dimensions. The guidelines were modelled on Statistics Canada's Quality Guidelines, which similarly provide non-exhaustive best practices for principles of data quality.

TBS and Statistics Canada will periodically review and update the Framework in consultation with federal partners in order to ensure its continued relevance and value to the GC. As well, TBS will advance efforts to operationalize the Framework by working to embed or reference it in TB policy instruments, governance processes and frameworks, and data sharing agreement templates.

The following federal organizations took part in the development of the Framework: Agriculture and Agri-Food Canada (AAFC); Canada Border Services Agency (CBSA); Canada Mortgage and Housing Corporation (CMHC); Canada Revenue Agency (CRA); Canada School of Public Service (CSPS); Canadian Food Inspection Agency (CFIA); Canadian Human Rights Commission (CHRC); Canadian Institutes of Health Research (CIHR); Canadian Nuclear Safety Commission (CNSC); Canadian Space Agency (CSA); Correctional Service Canada (CSC); Crown-Indigenous Relations and Northern Affairs Canada (CIRNAC); Department of Justice Canada (JUS); Department of National Defence (DND); Elections Canada; Employment and Social Development Canada (ESDC); Environment and Climate Change Canada (ECCC); Fisheries and Oceans Canada (DFO); Global Affairs Canada (GAC); Health Canada; Immigration, Refugees and Citizenship Canada (IRCC); Indigenous Services Canada (ISC); Innovation, Science and Economic Development Canada (ISED); Library and Archives Canada (LAC); Natural Resources Canada (NRCan); Privy Council Office (PCO); Public Health Agency of Canada (PHAC); Public Services and Procurement Canada (PSPC); Service Canada; Shared Services Canada (SSC); Standards Council of Canada (SCC); Statistics Canada; Transport Canada; Treasury Board of Canada Secretariat (TBS); and Veterans Affairs Canada (VAC).

Appendix D: References

Algorithmic Impact Assessment Tool: <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html>

Budget 2021 - A Recovery Plan for Jobs, Growth, and Resilience: <https://www.budget.gc.ca/2021/home-accueil-en.html>

Canada.ca Content Style Guide: <https://www.canada.ca/en/treasury-board-secretariat/services/government-communications/canada-content-style-guide.html>

Data Management Body of Knowledge, 2nd Edition (DMBOK2): DMBOK - Data Management Body of Knowledge (dama.org)

Digital Operations Strategic Plan: 2021-2024: <https://www.canada.ca/en/government/system/digital-government/government-canada-digital-operations-strategic-plans/digital-operations-strategic-plan-2021-2024.html>

Direction on the Secure Use of Commercial Cloud Services: <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/cloud-services/direction-secure-use-commercial-cloud-services-spin.html>

European Statistics Code of Practice for the National Statistical Authorities and Eurostat (EU Statistical Authority): <https://ec.europa.eu/eurostat/documents/4031688/8971242/KS-02-18-142-EN-N.pdf/e7f85f07-91db-4312-8118-f729c75878c7?t=1528447068000>

Government of Canada Core Subject Thesaurus: <https://canada.multites.net/cst/index.htm>

Government of Canada Digital Standards: <https://www.canada.ca/en/government/system/digital-government/government-canada-digital-standards.html>

Government of Canada Enterprise Architecture Framework: <https://www.canada.ca/en/government/system/digital-government/policies-standards/government-canada-enterprise-architecture-framework.html>

Government of Canada Security Control Profile for Cloud-based GC Services: <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/cloud-services/government-canada-security-control-profile-cloud-based-it-services.html>

Government of Canada Standards on APIs: <https://www.canada.ca/en/government/system/digital-government/modern-emerging-technologies/government-canada-standards-apis.html>

ISO 8000-2, Data quality — Part 2: Vocabulary: <https://www.iso.org/obp/ui/#iso:std:iso:8000:-2:ed-4:v1:en>

ISO 8000-8, Data quality — Part 8: Concepts and measuring: <https://www.iso.org/obp/ui/#iso:std:iso:8000:-8:ed-1:v1:en>

Library and Archives Canada Act: <https://laws-lois.justice.gc.ca/eng/acts/L-7.7/index.html>

Open Government Data and Information Quality Standards (Draft): <https://open.canada.ca/ckan/en/dataset/bfb87332-5da3-5780-9546-8722a389c91c> Privacy Act: <https://laws-lois.justice.gc.ca/ENG/ACTS/P-21/index.html>

Quality Dimensions, Core Values for OECD Statistics and Procedures for Planning and Evaluating Statistical Activities: <http://www.oecd.org/sdd/21687665.pdf>

Reid, Giles, Zabala, Felipa and Holmberg, Anders. "Extending TSE to Administrative Data: A Quality Framework and Case Studies from Stats NZ" *Journal of Official Statistics*, vol.33, no.2, 2017, pp.477-511. <https://doi.org/10.1515/jos-2017-0023>

Report to the Clerk of the Privy Council: A Data Strategy Roadmap for the Federal Public Service: <https://www.canada.ca/en/privy-council/corporate/clerk/publications/data-strategy.html>

Statistics Canada Necessity and Proportionality Framework: <https://www.statcan.gc.ca/en/trust/address>

Statistics Canada Quality Assurance Framework: <https://www150.statcan.gc.ca/n1/en/pub/12-586-x/12-586-x2017001-eng.pdf?st=hLNiTVy9>

Statistics Canada Quality Guidelines: <https://www150.statcan.gc.ca/n1/pub/12-539-x/2019001/ensuring-assurer-eng.htm>

Treasury Board Directive on Automated Decision-Making: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>

Treasury Board Directive on Open Government: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=28108>

Treasury Board Directive on Privacy Impact Assessment: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=18308>

Treasury Board Directive on Service and Digital: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32601>

Treasury Board Guidance for Drafters of Treasury Board Submissions: <https://www.canada.ca/en/treasury-board-secretariat/services/treasury-board-submissions/guidance-for-drafters-of-treasury-board-submissions.html>

Treasury Board Guideline on Service and Digital: <https://www.canada.ca/en/government/system/digital-government/guideline-service-digital.html>

Treasury Board Policy on Government Security: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=16578>

Treasury Board Directive on Security Management: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32611>

Treasury Board Policy on Service and Digital: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32603>

Treasury Board Standard on Geospatial Data: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=16553>

Treasury Board Standard on Metadata: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=18909>

UK Government – The Government Data Quality Framework: <https://www.gov.uk/government/publications/the-government-data-quality-framework>

United Nations National Quality Assurance Frameworks Manual for Official Statistics: <https://unstats.un.org/unsd/methodology/dataquality/references/1902216-UNNQAFManual-WEB.pdf>

[1] In this document, the term ‘user’ generally refers to a data consumer who needs high-quality data to support a policy, program, service, or other initiative in the federal government. The data can be used for the purpose for which it was initially obtained or reused for a consistent or other purpose, as permitted under privacy, security, and other applicable legislation. Users leverage the Government Data Quality Framework to identify, communicate, assess, report on, and help address issues of data quality in consultation with the appropriate stakeholders (e.g., data providers, data custodians, data policymakers, data stewards, data architects, subject matter experts, security and privacy officials).

This page was last edited on 16 December 2024, at 21:10.

Content is available under [Creative Commons Attribution 4.0 International License](#) unless otherwise noted.