Data Strategy for the Federal Public Service - Annexes

Français

The following is an evergreen list of terms and complementary strategies for the 2023-2026 Data Strategy for the Federal Public Service.

Contents

Glossary of Terms
- Data
- Aggregated data
- Disaggregated data
- Data flow
- Data governance
- Data management
- Data portability
- Data quality
- Data security
- Data standards
- Data steward
- Domain steward
- Data stewardship
- FAIR Data Principles
- Interoperability
- Privacy

Domain-Specific Strategies

References

Glossary of Terms

The following definitions are intended to support a common understanding of key terminology when reading the 2023-2026 Data Strategy for the Federal Public Service. They are intended to be a source of collaboration and knowledge sharing and are not official policy definitions.

Data

> The representation of information, in a manner suitable for storage, communication, interpretation, or processing by human beings or by automatic means, and from which knowledge can be drawn, including structured or unstructured forms. Often a set of values of subjects with respect to qualitative or quantitative variables representing facts, statistics, or items of information in a formalized manner.

- Statistical data refers to data used to produce official statistics (often from a census, survey statistical register or administrative source) by government agencies or other entities working on behalf of the government.
- Administrative data refers to data and information collected by organizations, government agencies or other public entities as a part of their ongoing operations. Examples include records of births and deaths, data collected by satellites, or records about the flow of goods and people across borders.[1][2][3][4][5][6][7]

Aggregated data

Unit level data that has been combined and summarized, often from multiple sources, into a collective form, often for the purposes of statistical analysis. Aggregate data allows for greater analysis and insight about particular groups based on specific variables, such as age or gender.[8][9]

Disaggregated data

Compiled or aggregate data that has been separated or broken down into smaller information units for the purposes of analysis. Disaggregated data allows for detailed analysis and insight about various subsets or outcomes within a larger data set. Data can be disaggregated by variables such as income or socio-cultural background.[8]

Data flow

The circulation or movement of computerised data and information through interoperable systems and across organizations, geopolitical regions or jurisdictions.[10]

Data governance

A system of decision rights and accountabilities, responsibilities and rules for the management of the availability, usability, integrity and security of the data and information to enable coherent implementation and co-ordination of data stewardship activities as well as increase the capacity (technical or otherwise) to better control the data value chain, and the resulting regulations, policies and frameworks that provide enforcement. This includes the systems within an enterprise, organization or government that define who has authority and control over data assets and how those data assets may be used, as well as the people, processes, tools and technologies required to manage and protect data assets .[1][11][12][13][14][15]

Data management

A discipline that directs and supports effective and efficient management of information and data in an organization or public administration, from planning and systems development to disposal or long-term preservation. Data management involves the development, execution, and supervision of plans, policies, practices, concepts, programs, and the accompanying range of systems that contribute to the organizational or governmental mandates and to public good, as well as the maintenance of data processes to meet ongoing information lifecycle needs. It enables the delivery, control, protection, and enhancement of the value of data and information assets through integrated, user-based approaches. Key components of data lifecycle management include a searchable data inventory, reference and master data management, and a quality assessment framework.[5][15][16][17][18]

Data portability

The capacity of digital data and information to be transmitted or circulated through interoperable applications or systems and across organizations or geopolitical regions. Data portability enables data subjects to have clear and manageable access to their personal data, which they have provided to a controller in a structured, commonly used, machine-readable and interoperable format, and are free to transfer it to another controller without undue burden.[19][20]

Data quality

The ‘quality’ of data refers to its fitness for purpose, often measured by such criteria offered in the bullet below. Data quality assurance measures are used to assess and improve the quality of data. Quality assurance measures planning, implementation, and control of activities that apply quality management techniques to data (whether statistical, administrative, or otherwise) and the statistical production process, to assure data is fit for purpose, which means that it is both usable and relevant in a primary or other use-context, and meets the needs of data users. Different users may have different needs that must be balanced.

- Many organizations – within Canada and internationally – have a set of criteria defining data quality. These often include concepts such as: *relevance*, *reliability, consistency*, *credibility*, *completeness*, *accuracy*, *timeliness*, *accessibility*, *comparability*, *interpretability, coherence*, and *proportionality*, which all contribute to the data and information’s overall quality and value.[7][16][21][22][23][24][25][26]

Data security

The definition, planning, development, and execution of security policies and procedures used to provide proper authentication, authorization, access, and auditing of data and information assets. Data security enables the protection of privacy, confidentiality, and integrity, as well as the maintenance of trust and social license to operate.[1][6][16][17]

Data standards

Data standards are the rules and specifications by which data are described, defined and recorded. In order to share, exchange, and understand data, standardized formats and meanings are needed. Examples of data standards include data models, reference data, identifier schemas, and statistical standards. The use of data standards enables the integration of data over time and across different data sources, as well as reduces the resource requirements associated with many aspects of survey development and maintenance. [1][14][27][28]

Data steward

The role(s) accountable for the management of data assets and resources from a strategic perspective. Data stewards are responsible for ensuring that the data acquisition, entry, quality, interoperability, and overall management supports organization's needs, while also ensuring adherence to social license, legislative, and regulatory requirements. They work with stakeholders and other deliberative or advisory bodies to develop definitions, standards and data controls, and perform key functions in the ideation and implementation of data policies that are scalable, sustainable, and significant. [1][14][29]

Domain steward

(Also called *domain lead*, *subject area steward*, *data domain steward*, or *business data steward*) A role within a data stewardship program, which is accountable for a particular data domain. The domain steward is the leader of the domain's stewardship team, will represent their domain on various data stewardship committees or data governance councils, and will help define, implement, and enforce data management policies and procedures within their specific Data Domain. Domain stewards are essential to a successful data governance program. Employing domain stewardship and domain data stewards is a way to govern data across functional areas of the enterprise. [13][30][31][32][33][34]

Data stewardship

Data stewardship is a discipline that directs and supports the ethical and responsible creation, collection, management, use, and reuse of data, and is applicable at all scales – from the national or data system level, to the organization or enterprise level, or to the individual or dataset. Data stewardship programs and processes are formalized through repeatable and automated business processes, established roles and accountabilities, and the use of metrics and audits in order to continuously improve data quality. Data stewardship operations influence proactive and responsible data practice to help deliver data strategies, maintain trust, and promote accountability, and it is enabled though good data governance and data management, which provide oversight of data assets throughout their lifecycle to ensure their proper care. [13][14][15][29]

FAIR Data Principles

Set of data principles, which define characteristics that modern data resources, tools, vocabularies and infrastructures should demonstrate to facilitate the discovery and reuse of data by other parties. FAIR stands for:

F - Findable and easily searchable
A - Accessible and easy to use
I - Interoperable and more easily interpretable
R - Re-usable data that is easy to share and use[35].

Interoperability

Interoperability is the ability to access, process and exchange data from multiple sources, then integrate that data for mapping, visualization, and other forms of meaningful representation and analysis. This allows systems and organizations to work together (inter-operate) towards mutually beneficial goals by sharing information and exchanging data. In order to be interoperable, data should follow established data standards to ensure that it is easily compared over time, across jurisdictions, and within and between departments. There are five key layers of interoperability:

1. Legal interoperability is about ensuring that organizations operating under different policies and strategies are able to work together.
2. Operational interoperability is about ensuring how organizations align their business processes, responsibilities and expectations to achieve mutual beneficial outcomes.
3. Semantic interoperability is about ensuring consistent meaning and optimal comparability of data with the use of conceptual models, vocabularies and ontologies.
4. Syntactic interoperability is about format. It allows us to explicitly define the common representations and exchange models.

5. System interoperability is about defining the infrastructure and communication protocols to be used during the exchange process.[5][15][36][37][38][39]

Privacy

Privacy describes the degree of protection and confidentiality that personal information and data will be accorded. For Canadian federal institutions, privacy requirements regulate the creation, collection, use, disclosure, protection, retention and disposal of personal information. Privacy can include guiding principles such as accountability, transparency, security, openness, and the rights to redress and to access one's own personal information.[1][15][16][40]

Domain-Specific Strategies

The following is a list of domain-specific strategies that complement the renewed Data Strategy for the federal public service. Additional strategies will be added as they are published.

Pan-Canadian Health Data Strategy (https://www.canada.ca/en/public-health/programs/pan-canadian-health-data-strategy.html) - the strategy aims to support the effective creation, exchange, and use of health data for the benefit of Canadians and the public health systems they rely on. A collaborative approach to develop and deliver the strategy is being taken through federal/provincial/territorial co-development of the strategy, which is informed by the latest research findings, public health and data experts, and an Expert Advisory Group to provide guidance as the work evolves.

The National Freshwater Data Strategy is under development to enable and improve access to data for safe, clean and well-managed freshwater. The Canada Water Agency is collaborating with partners across federal and provincial departments and agencies, Indigenous organizations and governments, and experts across the country.

References

1. Organization for Economic Co-operation and Development (2008). *OECD Glossary of Statistical Terms*, OECD Publishing, Paris. https://doi.org/10.1787/9789264055087-en.
2. Organisation for Economic Co-operation and Development (2021). Recommendation of the Council on Enhancing Access to and Sharing of Data. OECD Legal Instruments. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463
3. Statistics Canada (2016). Statistics Canada Policy on the Use of Administrative Data Obtained under the Statistics Act. Ottawa, ON: Her Majesty the Queen in Right of Canada. https://www.statcan.gc.ca/en/about/policy/admin_data
4. Statistics Canada (2023). Administrative Data. Statistics Canada. https://www.statcan.gc.ca/en/our-data/where/administrative-data
5. Government of Canada, Treasury Board Secretariat (2019). *Policy on Service and Digital*. Ottawa, ON: Her Majesty the Queen in Right of Canada. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32603
6. Economic Commission for Europe of the United Nations (UNECE). (2000). Terminology on Statistical Metadata in *Conference of European Statisticians Statistical Standards and Studies*. (53), Geneva. https://digitallibrary.un.org/record/442455
7. United Nations Departments of Economic and Social Affairs (2019). *United Nations National Quality Assurance Frameworks Manual for Official Statistics* [PDF]. https://unstats.un.org/unsd/methodology/dataquality/references/1902216-UNNQAFManual-WEB.pdf

8. National Collaborating Centre for Indigenous Health (2010). The Importance of Disaggregated Data. https://www.nccih.ca/docs/context/FS-ImportanceDisaggregatedData-EN.pdf
9. Statistical Data and Metadata eXchange (2020). SDMX Glossary Version 2.1. https://sdmx.org/wp-content/uploads/SDMX_Glossary_version_2_1-Final-2.docx
10. Organisation for Economic Co-operation and Development (1985). *Declaration on Transborder Data Flows*. OECD: Better Policies for Better Lives. https://www.oecd.org/sti/ieconomy/declarationontransborderdataflows.htm
11. Data Governance Institute (n.d.). *Governance and Decision Making*. Data Governance Institute. https://datagovernance.com/governance-and-decision-making/
12. Organisation for Economic Co-operation and Development (2019). Data Governance in the Public Sector *In The Path to Becoming a Data-Driven Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris. https://doi.org/10.1787/059814a7-en.
13. Plotkin, D. (2021). Data Stewardship: An Actionable Guide to Effective Data Management and Data Governance (2nd Ed.). London, UK: Academic Press.
14. Statistics Canada (2021). *Enterprise Information and Data Management Glossary* [PDF]. Unpublished internal departmental document.
15. Statistics Canada (2019). Statistics Canada Data Strategy: Delivering insight through data for a better Canada https://www.statcan.gc.ca/en/about/datastrategy
16. Statistics Canada (2021). *Statistics Canada's Aspects of Data Stewardship* [PDF]. Unpublished internal departmental document.
17. Data Management Association (DAMA) (2017). DAMA-DMBOK: Data Management Body of Knowledge (2nd Ed.). Basking Ridge, NJ: Technics Publications.
18. Statistics Canada (2020). *Data Literacy Competencies*. Statistics Canada. https://www.statcan.gc.ca/en/wtc/data-literacy/compentencies
19. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] Official Journal of the European Union, Legislation Series 119. https://eur-lex.europa.eu/eli/reg/2016/679/oj
20. Government of Canada, Innovation, Science and Economic Development Canada (2019). *Canada's Digital Charter in Action: A Plan by Canadians, for Canadians.* Ottawa, ON: Her Majesty the Queen in Right of Canada. https://ised-isde.canada.ca/site/innovation-better-canada/en/canadas-digital-charter/canadas-digital-and-data-strategy
21. European Commission, Eurostat (2003). *Assessment of quality in statistics - Definition of Quality in Statistics*, Working Group, Luxembourg, October 2003. https://ec.europa.eu/eurostat/documents/64157/4373735/02-ESS-quality-definition.pdf
22. European Commission, Eurostat (2020). Quality assurance framework of the European statistical system: version 2.0, Publications Office, 2020. https://data.europa.eu/doi/10.2785/847733
23. Government of Canada (2022). *GC Data Quality Framework*.https://wiki.gccollab.ca/GC_Data_Quality_Framework#Background
24. Organisation for Economic Co-operation and Development (2002). Measuring the Non-Observed Economy: A Handbook. Paris, France: OECD Publications. https://www.oecd.org/sdd/na/measuringthenon-observedeconomy-ahandbook.htm
25. Statistics Canada (2002). *Statistics Canada's Quality Assurance Framework*. Ottawa, ON: Minister of Industry. https://www150.statcan.gc.ca/n1/en/pub/12-586-x/12-586-x2002001-eng.pdf?st=QDz6ld3y
26. Wang, R.Y. and Strong, D.M. (1996) *Beyond Accuracy: What Data Quality Means to Data Consumers*. Journal of Management Information Systems, 12, 5-33. https://doi.org/10.1080/07421222.1996.11518099
27. International Organization for Standardization (2016). Data quality — Part 61: Data quality management: Process reference model (ISO standard no. 8000-61:2016) https://www.iso.org/obp/ui/#iso:std:iso:8000:-61:ed-1:v1:en

28. Standards Council of Canada (2020). *What are standards?* Standards Council of Canada. https://www.scc.ca/en/standards/what-are-standards
29. Organisation for Economic Co-operation and Development (2018). Governing open data for sustainable results, in Open Government Data Report: Enhancing Policy Maturity for Sustainable Impact, OECD Publishing, Paris. https://read.oecd-ilibrary.org/governance/open-government-data-report/governing-open-data-for-sustainable-results_9789264305847-4-en#page1
30. Marco, D.P. (n.d.). Data Stewardship Roles: A Complete Guide. DataManagementU. https://www.ewsolutions.com/data-stewardship-roles-a-complete-guide/
31. Seiner, R.S. (2007). The Data Stewardship Approach to Data Governance: Chapter 7. The Data Administration Newsletter. https://tdan.com/the-data-stewardship-approach-to-data-governance-chapter-7/6173.
32. Loshin, D. (2001). Enterprise Knowledge Management: The Data Quality Approach. The Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann Publishers.
33. Strengholt, P. (2021). Data Domains and Data Products. Towards Data Science. https://towardsdatascience.com/data-domains-and-data-products-64cc9d28283e
34. University of Washington (2023). Data Stewardship – UW's Data Domains and Councils. University of Washington data Governance. https://datagov.uw.edu/data-stewardship/
35. Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al*. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data 3*, 160018. https://www.nature.com/articles/sdata201618
36. European Commission (2017a). European Political Strategy Centre, Enter the data economy: EU policies for a thriving data ecosystem. Publications Office 21:11. https://data.europa.eu/doi/10.2872/33746
37. European Commission (2017b). European Interoperability Framework. Luxembourg: Publications Office of the European Union. https://ec.europa.eu/isa2/sites/default/files/eif_brochure_final.pdf
38. Data Documentation Initiative Alliance (2021). *DDI Alliance Glossary*. DDI Alliance. https://ddialliance.org/resources/ddi-glossary
39. Chapurlat, V., Daclin N. (2012). System interoperability: definition and proposition of interface model in MBSE Context. IFAC Proceedings Volumes, 45(6), 1523-1528. https://www.sciencedirect.com/science/article/pii/S1474667016333675
40. Government of Canada, Treasury Board Secretariat (2019). Directive on Privacy Practices. Ottawa, ON: Her Majesty the Queen in Right of Canada. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=18309

Retrieved from "https://wiki.gccollab.ca/index.php?title=Data_Strategy_for_the_Federal_Public_Service_-_Annexes&oldid=121971"

This page was last edited on 13 June 2024, at 12:39.

Content is available underCreative Commons Attribution 4.0 International Licenseunless otherwise noted.