NDSF 2020
Kanata Declaration

Version - August 25, 2020

Authored by Attendees of the 2019 and 2020 NDSF Summits and Members of the Canadian RDM Ecosystem

DOI: 10.5281/zenodo.4002076

Declaration

The Kanata Declaration[1] reflects a broad consensus from the Research Data Management (RDM) community on priorities for the data management (DM) component of the emerging Digital Research Infrastructure (DRI) Strategy. The intended targets are first and foremost Canadian researchers, but also the New DRI Organization (NDRIO), as well as decision makers and participants in the broader research community.

Recognize that researchers are the focal point for this conversation, and their needs are central to the design and delivery of services and resources both across and within specific domains; that creative, researcher-centric engagement, with agency and autonomy to share their data, together with support for participation from all citizens, enriches the open science[2] community, and creates an environment where all research can be questioned and advanced.

Recognize that the concept of open science means different things to different people, but ultimately encompasses: the *Philosophy* of open science which is the concept of "advancing research on the shoulders of giants", a framing of doing research that is widely recognized; the *Policy* of open science which encourages or mandates the sharing of data produced with public funds, or via a journal publisher's requirements; and the *Practice* of open science which is where every discipline defines its own best practices and tools, including whether or not it is beneficial to share research outputs, and when, how, and what to share.

Recognize the public good of data as a foundation for knowledge and evidence-based decision making in Canadian and international research and society, and its impact on the quality of life and wellbeing of all Canadians.

Recognize the imperative to innovate and lead in this space to ensure the future viability, integrity, and stability of Canadian data assets, their availability for reuse in Canada and internationally, and their impact on research excellence.

Recognize the importance of a supranational approach to international engagement and Canadian leadership in this context, as a prerequisite for building capacity and confidence in research, and that this engagement needs to be effectively resourced.

Recognize that research is becoming more data intensive; that data are an important asset that should be effectively managed with attention to ensuring (meta)data[3] quality via domain-specific practices, and using trusted data repositories, to preserve and enhance their value; that impact and innovation come from ground level initiatives, supported with resources, expertise and

---

[1] This is a major revision of the [2019 Kanata Declaration](#) which emerged from the 2019 NDSF Summit. Key updates include the incorporation of feedback from the 2020 NDSF Summit Breakout sessions, collaborative notes kept during the Summit, and feedback on each of the original Declaration Statements that were solicited via a survey tool from February to April 2020.

[2] One definition of open science: *Open Science is the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods.* (FOSTER: www.fosteropenscience.eu/taxonomy/term/100)

[3] (*meta)data* is a convention intended to apply to both metadata (data about data) and the data assets themselves.

training, at local, regional, and national layers, to leverage common objectives, approaches, systems, knowledge, and tools.

Recognize that to effectively steward these data assets, Canadian funders, government agencies, universities and other infrastructure providers need to ensure the sustainable long-term support for research platforms, data repositories (including certification like CoreTrustSeal[4]), and associated tools and services, with particular attention to international best practices, such as FAIR[5] and TRUST[6] principles.

Recognize that Research Data Management (RDM) and Digital Research Infrastructure (DRI) are not just about stable national and global infrastructure, but fundamentally about people and institutions that require ongoing funding support, as well as a responsive process and policy framework to encourage and support best research practices at all layers and jurisdictions.

Highlight some of the key gaps and opportunities identified from the National Data Services Framework (NDSF) discussions.[7]
1. Architecture: the architecture necessary to create a federated system of DM platforms, tools and resources that ensure interoperability and access to all Canadian researchers.
   a. Support for a curated system that facilitates discovery and replication of data from disparate platforms, along with clear ownership and provenance information, into new data assets and research activity.
   b. Diagrams and descriptions of an agreed upon architecture showing national and local levels, and showing all platforms, tools, and services.
2. Data: the common standards and tools that support the FAIR principles across disciplines, while ensuring access to rich domain-specific (meta)data.
   a. Document and encourage best practices for data review and quality assurance as an integral part of the research process in understanding existing data.
   b. Agreement on a minimum common denominator for a metadata description that facilitates mapping from existing metadata descriptions, leading to better cross-domain discovery and access, and that is based on metadata standards such as schema.org, ISO 19115, and the DataCite Metadata Schema, or on an appropriate linked data model with support for domain-specific metadata.
3. Services: generic and discipline-specific services directly supporting researchers and institutions, as well as funder, publisher, and institutional policy.
   a. Sustainable support (e.g. financial, training, shared infrastructure) for service providers, including data scientists, data stewards/curators, librarians, data management specialists, IT specialists, repository managers, and research officers to support researchers from all disciplines.

---

[4] https://www.coretrustseal.org/

[5] FAIR stands for Findable, Accessible, Interoperable, Reusable. The FAIR Guiding Principles for scientific data management and stewardship. Mark D. Wilkinson et al. Scientific Data. March 2016.

[6] TRUST stands for Transparency, Responsibility, User focus, Sustainability, and Technology. The TRUST Principles for digital repositories. Dawei Lin et al. Scientific Data. May 2020.

[7] Statements are made in the context of the 6-part EOSC model, which was used in the 2019 and 2020 Summits.

b. Support for a sustainable framework of data stewardship education and training directly to researchers (e.g. iKT framework[8]), without funding barriers, ensuring they maintain autonomy in their respective disciplines, but at the same time have access to, and are aware of, best practices, tools, and services.

c. Commitment to user-centered design for practices, tools, and services, recognizing that time and resources invested into usable, appreciated infrastructure pays dividends in productivity and adoption.

d. Sustainable support (e.g. financial, training, shared infrastructure) for existing and emerging platforms and services, including institutional and domain-specific repositories.

e. Certification of institutions, platforms, and staff at the national level, using international best practices, such as CoreTrustSeal, ISO 16363, etc.

4. Access & Interfaces: human and machine interfaces that make it simpler to deposit and access data, including robust support for privacy and security where appropriate, Indigenous-specific data and traditional knowledge, as well as multiple languages.

a. Support for a comprehensive index of RDM services and resources at local, regional, and national levels, including links to appropriate international efforts, and the ability for its integration into local systems.

b. Support for centralized identity management via integration of identity/authorization across multiple services, paired with robust cybersecurity and effective bi-directional (e.g. pull/push of metadata) support for the ORCID registry[9].

c. Promotion of best practices for data licensing, policies, and usage restrictions, including a preference for open data where appropriate.

5. Rules: policy and process scaffolding that facilitates participation by all actors in the research ecosystem, supports jurisdictional contexts, and engenders trust.

a. Encourage participation of early-career researchers (ECRs: graduate students, post-grads, etc.), including filling the gap in accessibility to funding, as well as recognition for promotion and tenure, for community building, data sharing, and secondary data use as equally important to creating new datasets.

b. Clarify and develop support (including financial) for researchers to respond to RDM policies in all national and provincial research funding programs.

6. Governance: the framework that ensures representation by all actors in the development and sustainability of Canada's digital research infrastructure and ecosystem.

a. Promote culture change in the research community via an incentive model that facilitates the adoption of RDM best practices, participation by all researchers in all disciplines, as well as by institutional RDM support services, and recognition of data citations and reuse.

b. Support the development of institutional RDM strategies that respond to funder policies and intersect effectively with institutional data governance models.

c. Provide funding support for platform developers to add data deposit and metadata enrichment services to their systems, including the management of all

---

[8] Institute for Knowledge Translation. https://www.knowledgetranslation.org/
[9] https://orcid.org/about/what-is-orcid/mission

appropriate persistent identifiers (PIDs), and the submission of metadata about deposited data into the ORCID registry.

d. Promote the integration of a cross-sectoral approach, including higher education, non-profit, commercial, and provincial/federal governments, to the governance bodies of all NDSs.

Resolve to raise awareness about, and ensure concurrent growth occurs in, capacity for RDM infrastructure and services, along with culture change in researcher adoption of RDM services, and especially with ECRs.

Resolve to ensure continuous dialogue and engagement within and across domains and functional teams, using a wide range of tools and avenues, among the diverse community of researchers, funders, users, service providers, and other actors (e.g. patient groups, Indigenous communities), to build trust, understanding, and consensus around Canadian RDM and DRI.

Resolve to establish and monitor an integrated national DRI ecosystem that includes the RDM community, that sustains, enhances, and advances Canada's capacity and impact in the global research and innovation community, while continuing to support existing infrastructures in specialized areas.

Resolve to identify and articulate sustainable ongoing annual support for core RDM infrastructure (e.g. tools, platforms, persistent identifiers and other services) in the DRI ecosystem.

Resolve to ensure that First Nations, Métis, Inuit and broader global Indigenous communities, and Indigenous-led considerations of traditional knowledge, are equal participants in the conversation for RDM and DRI, and an integral part of all activities in the ecosystem. This includes:
1. consideration of Indigenous data and Traditional Knowledge (TK);
2. use of Indigenous-specific approaches such as the OCAP[10] and CARE[11] Principles;
3. use of and mechanisms for clarifying Indigenous-specific language regarding use and access, such as the TK and Biocultural (BC) Labels;
4. development of specific metadata fields that recognize Indigenous interests, supporting better provenance practices for Indigenous data, and encouraging use by researchers of new notification tools (e.g. TK and BC Notices), participatory meetings and gatherings.

Resolve to put into place appropriate privacy and security safeguards for all data, including appropriate support for qualitative, sensitive and confidential data, involvement of ethics boards or equivalent bodies, training on the ethical use of data, implementation of cybersecurity audits, and ensuring appropriate consent and/or access control for various forms of sensitive data (e.g. data that must be protected against unwanted disclosure).

---

[10] OCAP stands for Ownership, Control, Access, and Possession. The First Nations Principles of OCAP. https://fnigc.ca/ocap

[11] CARE stands for Collective benefit, Authority to control, Responsibility and Ethics. The CARE Principles for Indigenous Data Governance. https://www.gida-global.org/care

Commit to advancing FAIR data and open science through the establishment of a national data repository network defined by cross-repository discovery, a core set of common standards, trusted certification of core repositories, and including a commitment to key principles-based frameworks (i.e. FAIR, TRUST, CARE), and long-term data stewardship.

Commit to developing national support and funding for RDM training in all domains throughout the entire research life cycle, across all components of the ecosystem, including scalable online options, and incorporating elements regarding the control and ownership of data, data sovereignty, including Indigenous-specific research data, and sensitive data.

Commit to a service model that delivers data and services to all Canadian researchers, regardless of location, bandwidth limits, or access to resources.

Commit to supporting a Canadian network of data champions from all domains to ensure that proper support, curation, and stewardship services are available to all researchers in all institutions regardless of size.

Commit to ensuring that the control and ownership of data remains with researchers and/or Indigenous communities, and/or is made clear in data availability statements, and that policy makers, funders, and service providers are aware of key concerns, and actively working to respond to these concerns.

Commit to continuing to establish Canada as a leader internationally, working collaboratively with the international community, adopting best practices and standards where appropriate, in further development of the RDM ecosystem.

Call researchers and RDM service providers to join as partners to share and transfer the knowledge required for exemplary stewardship of research data.

Call together representatives from key stakeholder organizations and researchers to develop and maintain a dynamic high-level RDM Roadmap to help prioritize investment and development of DRI for RDM.[12]

Commit to reporting back to NDSF participants and the broader stakeholder community on the developments of the strategy for support for RDM in Canada.

Supporting Materials

- RDC Zenodo Community
  - Includes all published material from the last 3 NDSF Summits, as well as other RDC publications.
- Research Data Management in Canada: A Backgrounder. [EN | FR]
- Data Management Roadmap - 2019-2024.

---

[12] A foundation for this effort can be found in the *Data Management Roadmap: 2019-2024*.