



# Canadian National Data Services Framework: Discussion Document

Prepared for the *2nd National Data Services Framework Summit*  
(January 24-25, Ottawa, ON)

Research Data Canada

*NDSF Working Group,  
Infrastructure Committee*

*Developed by the RDC stakeholder community, including:  
David Baker, Eugene Barsky, Jane Burpee, Mark Leggott, Jeff Moon, Michael Sinatra, Byron Spencer, Laura Gerlitz*

Version 1.0

doi: 10.5281/zenodo.2536558

January 9, 2019

## Table of Contents

<b>Background and Definitions</b>	<b>3</b>
Definitions	3
<b>International Context</b>	<b>5</b>
<b>The Canadian Landscape</b>	<b>6</b>
Ecosystem Gaps and Opportunities	7
2018 CANARIE DM Funding	9
<b>Evolution of a Canadian NDSF</b>	<b>11</b>
<b>Next Steps</b>	<b>13</b>
<b>Resources</b>	<b>14</b>
<b>Appendix A: The Vienna Declaration on the European Open Science Cloud</b>	<b>15</b>
<b>Appendix B: NDSF Diagram</b>	<b>17</b>

# Background and Definitions

Research Data Canada (RDC) began using the term National Data Services Framework (NDSF) in 2016, primarily as a way to continue a broad stakeholder conversation focused on national digital research infrastructure and associated data management (DM). This was informed by discussions at previous national data summits<sup>1,2,3</sup>, and the Research Data Alliance (RDA) National Data Services Interest Group (NDSIG)<sup>4</sup>, which is co-chaired by Mark Leggott, RDC's Executive Director, Christine Kirkpatrick from the U.S. NDS, and Adrian Burton, from the Australian Data Commons. This document is intended as a backgrounder to the 2nd NDSF Summit<sup>5</sup>, and as a general discussion paper for the Canadian data management community.

## Definitions

The term National Data Services Framework was defined at the 1st NDSF Summit at 3 levels with respect to a Canadian approach to developing a more integrated and sustainable suite of national data services:

1. a conversation with all stakeholders at all levels;
2. agreement on best-practices, standards and protocols;
3. a suite of interoperable services and resources.

This 3-part definition was intended to facilitate a national conversation about how to best support stakeholders in the research ecosystem, and to encourage the cultural and policy changes needed to make research data open and Findable, Accessible, Interoperable, and Reusable (FAIR)<sup>6</sup>.

Subsequent conversations led to the development of a more formal definition, which has been proposed for inclusion in the *CASRAI Dictionary*<sup>7</sup>.

*Short Definition: A service that provides one or more data-related functions to applicable stakeholders and disciplines in a specific national context.*

*Extended Definition: A service that provides one or more data-related functions to applicable stakeholders and disciplines in a specific national context. Such services are often provided via a national partnership of data providers, data aggregators, community-specific groups, publishers, and cyber-infrastructure providers. In the context of research data functions such a data service would aspire to meet certain criteria such as: supporting/responding to the FAIR Data Principles and being appropriately*

---

<sup>1</sup> <https://www.rdc-drc.ca/download/background-to-the-2011-canadian-research-data-summit/?wpdmml=673>

<sup>2</sup> <https://www.rdc-drc.ca/download/report-of-the-canadian-research-data-summit/?wpdmml=675>

<sup>3</sup> <https://digitalleadership.ca/summit-2017/>

<sup>4</sup> <https://www.rd-alliance.org/groups/national-data-services.html>

<sup>5</sup> <https://zenodo.org/record/1035843>

<sup>6</sup> <https://www.nature.com/articles/sdata201618>

<sup>7</sup> [https://dictionary.casrai.org/Category:Research\\_Data\\_Domain](https://dictionary.casrai.org/Category:Research_Data_Domain)

*accessible to anyone in the said national context who receives or distributes public research funding. In the context of administrative data functions such a service would also aspire to meet certain criteria specific to that data domain. Such services would typically build on the data stewardship and sharing efforts underway within the said national context and would appropriately link with other such services - including international services. Such services will often be bundled into a framework, or supporting structure, based on standards, operations, policy, or more likely a combination of these, that promotes and delivers national data services in a coordinated fashion.*<sup>8</sup>

This definition became the basis of a group editing exercise for the NDSIG at the Berlin RDA Plenary. There were few edits to the text of the definition, and much of the discussion revolved around the use of the term “National”. As one might expect, this is a word that is both beneficial and challenging in facilitating conversations in any context where a political or geographic jurisdiction is considered. At the end of the conversation those present agreed that the use of “National” was not perfect, but was appropriate in framing the conversation in an international (and national) context.

Digital Research Infrastructure (DRI) is the preferred term in Canada for the specific services and resources needed to support the research enterprise. The 2015 Canada Foundation for Innovation (CFI) report “*Developing a digital research infrastructure strategy for Canada: The CFI perspective*”, defined it as:

*...the components that are collectively managed and operated as shared facilities and services for research institutions and users across the country because they are so large in scale, complexity and cost that they cannot be offered by a single institution. These components comprise computational capacity, data storage, technical services, research software, middleware, high-speed optical networks and research data management capabilities.*<sup>9</sup>

Research Infrastructure (RI) is a broader term, and also one more frequently in use in the European context than DRI.

*...facilities, resources and related services that are used by the scientific community to conduct top-level research in their respective fields and covers major scientific equipment or sets of instruments; knowledge-based resources such as collections, archives or structures for scientific information; enabling Information and Communications Technology-based infrastructures such as Grid, computing, software and communication, or any other entity of a unique nature essential to achieve excellence in research. Such infrastructures may be “single-sited” or “distributed” (an organised network of resources).*<sup>10</sup>

---

<sup>8</sup> <https://forum.casrai.org/t/national-data-service/939>

<sup>9</sup> <https://www.innovation.ca/sites/default/files/Funds/cyber/developing-dri-strategy-canada-en.pdf>

<sup>10</sup> [https://ec.europa.eu/research/infrastructures/pdf/eric\\_en.pdf](https://ec.europa.eu/research/infrastructures/pdf/eric_en.pdf)

The Research Data Lifecycle (RDL) is a common visual tool in the research data management (RDM) community, and is used to describe the basic steps in most research projects, as well as associated services and resources available to researchers at each stage. The diagram below was developed by the Leadership Council for Digital Research Infrastructure (LCDRI)<sup>11</sup>, and is intended to show the stages (outer circles) and associated service considerations (inner circles).

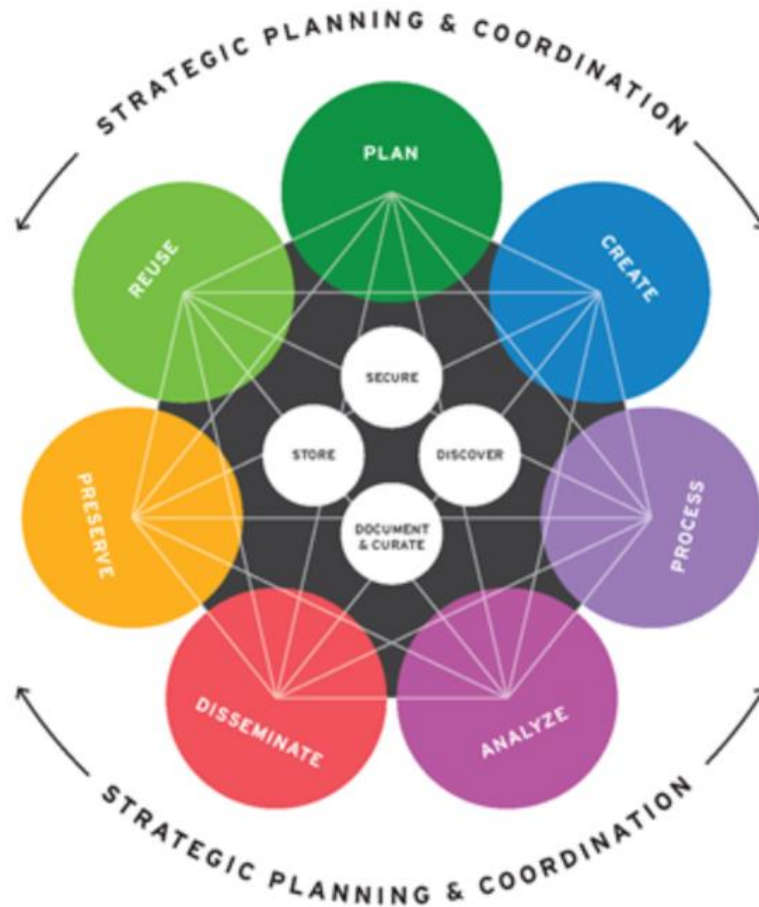


Figure 2. Data-Related Activities During the Research Process

## International Context

Many countries are actively working on the development of National Data Services: some have been in development for decades, and others are in the early stages. A recent report by the NDSIG<sup>12</sup> describes a number of these national efforts. Some key conclusions from these efforts include:

<sup>11</sup> <https://digitalleadership.ca>

<sup>12</sup> <https://drive.google.com/open?id=17iUyJ2icY0gFzMZGPWJyY5E0tUoukAtI4BFeronefv4>

1. the wide diversity of NDSs in a national context;
2. the lack of long-term funding/sustainability and associated business models in most of the cases;
3. the ongoing effort to define the details of NDS, which is an important part of developing an interoperable “open science framework” across the globe.

While there are a number of specific international examples that Canada can draw from, the European Open Science Cloud (EOSC)<sup>13</sup> could be the most aspirational. The core principles and commitments underlying the EOSC are most succinctly described in the *EOSC Declaration*.<sup>14</sup> Most recently, the *Vienna Declaration on the European Open Science Cloud* was made at a launch event for the EOSC on November 23rd.<sup>15</sup> This project is an instructive example for Canada, given that it:

1. assumes a federated model that builds on existing as well as emerging services and resources from member nations;
2. is working towards a federated governance model that reflects and respects individual national contexts;
3. has the FAIR principles as a foundational construct running throughout the effort.

Recent efforts in Australia to merge the efforts of 3 previously separate DRI organizations into the Australian Research Data Commons (ARDC), also provides an interesting history and set of use cases for efforts in Canada.<sup>16,17</sup> The Australian example is especially interesting as it shows the evolution of a suite of national organizations into a single, more cohesive approach to a sustainable DRI, something that was alluded to in Canada’s 2018 federal budget.

## The Canadian Landscape

The Canadian community has been discussing the idea of a national approach to data management for close to a decade, especially via the national meetings and Summits previously referenced, as well as the reports from CFI and the LCDRI. The LCDRI authored 3 documents (on *Advanced Research Computing*, *Data Management*, and *Coordination*) for Industry, Science and Economic Development (ISED) that made recommendations for proceeding with the development of a robust national approach to DRI. These reports were considered in the federal budgetary process, which announced an injection of \$572.5 million over 5 years to develop DRI, as well as additional annual funding. The details for how this effort will proceed are still being discussed. The submitted DM document is also currently being reworked into an *Environmental Scan of the DM ecosystem in Canada*, and will be available in 2019. The long-term goal is to update this document regularly to provide a dynamic reflection of DM efforts in Canada.

---

<sup>13</sup> <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

<sup>14</sup> [https://ec.europa.eu/research/openscience/pdf/eosc\\_declaration.pdf#view=fit&pagemode=none](https://ec.europa.eu/research/openscience/pdf/eosc_declaration.pdf#view=fit&pagemode=none)

<sup>15</sup> See Appendix A.

<sup>16</sup> [https://ardc.edu.au/wp-content/uploads/2018/10/ARDC\\_Strategic\\_Overview.pdf](https://ardc.edu.au/wp-content/uploads/2018/10/ARDC_Strategic_Overview.pdf)

<sup>17</sup> <https://ardc.edu.au/strategic-planning/capital-implementation-plan>

Against this backdrop, the Canadian stakeholder community has continued to facilitate a national conversation around the data management side of the DRI landscape, including via the upcoming 2nd NDSF Summit, to be held in January 2019.<sup>18</sup> The goal of this Summit, and related discussions and meetings, is to develop a coordinated and common view of a sustainable Canadian DM ecosystem going forward. This effort is bolstered by ongoing funding efforts (e.g. under existing CANARIE<sup>19</sup> and CFI<sup>20</sup> programs), national coordination/facilitation (e.g. RDC, CARL Portage), institutional/consortial activities (e.g. Cybera, Ontario Library Research Cloud, Compute Canada), development of national data repository options (e.g. Dataverse North, FRDR), development of national RDM platforms (e.g. DMP Assistant), development of training tools and resources (e.g. online modules from CIHR and Portage), and the continued development of research platforms and domain repositories (e.g. Ocean Networks Canada, CBRAIN, Polar Data Catalogue), all of which seek to provide enhanced DRI for Canadian researchers.

## Ecosystem Gaps and Opportunities

One of the most recent effort to identify gaps and opportunities in the DM landscape was the DM Community Consultation conducted by CANARIE in June 2018, as a way to identify funding priorities for their first DM funding call. The Consultation resulted in the development of 8 themes representing current gaps in the Canadian DM ecosystem.<sup>21</sup> It is important to understand that these illustrate the thinking in the Canadian context in mid-2018, and do not reflect all gaps and opportunities in the DM landscape.

1. Enriching (meta)data<sup>22</sup> and Discovery
  - a. As the diversity of research outputs increases, the challenge of finding specific researchers and resources also increases, especially at the national level. There is an opportunity to aggregate existing (meta)data from all systems and stages of the research lifecycle, as well as innovative approaches to enriching metadata (e.g. text mining, linking, entity extraction), that will improve the discovery of Canadian research outputs via an NDSF.
2. Federated Repositories / Interoperability
  - a. There are in the order of 200 Canadian “data repositories”, and likely many times this number of databases that contain useful research (meta)data. Each repository tends to present a unique architecture and (meta)data profile, and few are accessible via well-defined application program interfaces (APIs). There is an opportunity to use international standards and best practices to better document and make accessible Canada’s (meta)data repositories, and to develop smarter

---

<sup>18</sup> <https://www.rdc-drc.ca/activities/ndsf/>

<sup>19</sup> <https://www.canarie.ca/DM/>, <https://www.canarie.ca/software/>

<sup>20</sup> <https://www.innovation.ca/funded-projects>

<sup>21</sup> <https://www.canarie.ca/?wpdmdl=14726>

<sup>22</sup> (meta)data is a convenience used to refer to both metadata and data at the same time.

approaches to harvesting and linking these sources at the national level, thereby facilitating interdisciplinary research, discovery, and re-use.

3. Domain-Specific Repositories
  - a. Domain-specific repositories provide researchers with sources of (meta)data deposit and discovery, facilitating a response to funder/publisher mandates, and support domain-specific features that make (meta)data more useful. There is an opportunity to enhance Canada's domain repositories to ensure that they are interoperable with each other at the national level, and that these systems use internationally recognized best practices.
4. Data Deposit and Curation
  - a. Publishers and funders are adopting policies that require researchers to make data from their published outputs more accessible and reusable, but the options available to the researcher can be confusing and complex. There is an opportunity to create software services to guide and document the best approach to data management for each researcher, in a standards-based context with a minimum of effort (e.g. with the assistance of machine-mediated services), and according to practices that make their data FAIR.
5. Preservation
  - a. While it can be a challenge to determine what research (meta)data should be preserved and for how long, it is clear that (meta)data is not accessible and reusable over the long term unless it is preserved. There is an opportunity to build IT tools that facilitate this decision-making process in such a way that leverages appropriate best practices that ensure research outputs are accessible in a usable way for a reasonable period of time.
6. Persistent IDs / Citability
  - a. A key requirement for making research outputs accessible is the adoption of accepted best practices for unique IDs, whether for the researcher and their team, the variety of outputs, or the research equipment and services used. There is an opportunity to integrate Canadian research platforms with best practice standards for persistent IDs that make it easier to link the various outputs throughout the lifecycle, and attribute outputs to all research participants.
7. Data Access and Analytics
  - a. Researchers have increasingly diverse approaches to analyzing and transforming their data, whether in local systems or in the context of large national and international high-performance computing frameworks. There is an opportunity to build services that bring disparate sources of data to the researcher to facilitate data analysis at all stages of the research lifecycle.
8. Data Privacy and Security
  - a. In some disciplines (e.g. human health, biodiversity, social sciences), (meta)data cannot be easily shared without extensive and lengthy anonymization or protection processes, and in some cases cannot be shared at all. There is an opportunity to develop software services that facilitate this effort through semi-automated anonymization, as well as approaches that link datasets about individuals without unauthorized disclosure of personal information.



One example of an important issue not highlighted with this Consultation is that of Indigenous data, which requires special treatment as the property of the community about which it speaks. The *OCAP Principles* (Ownership, Access, Control, Possession)<sup>23</sup>, and efforts of the *International Indigenous Data Sovereignty Interest Group* of the RDA, are important in framing the discussion of Indigenous data in the Canadian context.<sup>24</sup>

## 2018 CANARIE DM Funding

The subsequent CANARIE DM Call resulted in the selection of 9 projects that addressed one or more of the above themes, as well as the FAIR principles and the NDSF. These projects were announced in November 2018, and include national DRI for generic use (i.e. accessible to any researcher in Canada) as well as domain-specific use (i.e. intended for use by any researchers in a specific community of practice). The projects will also be detailed in a map that will show intersections, areas of interoperability, overlap in software development, and other relevant information. Members of each of these funded projects will meet in Ottawa on January 23rd to discuss their efforts, and will also attend the 2nd NDSF Summit.

1. **Canadian Health Omics Repository, Distributed (CanDIG CHORD)** – *Led by Dr. Guillaume Bourque, McGill University*
  - a. CanDIG is a national project that allows collaborative analysis of human health genomics data distributed across the country, enabling stewards of this data complete, auditable control over data access. The CHORD project will create a federated Canadian national data service for privacy-sensitive genomic and related health data. It will also broaden the Canadian health research community's access to the technologies and services being built by CanDIG and its international partners in the Global Alliance for Genomics and Health.
2. **Dataverse for the Canadian Research Community** – *Led by Kate Davis, University of Toronto*
  - a. Dataverse is an open-source research data repository platform, developed by Harvard University's Institute for Quantitative Social Science with adopters and contributors from Canada, the US, and Europe. Originally architected to serve the needs of social science researchers with small to medium size data files, this project will adapt Dataverse's software architecture to address the needs of a broad range of researchers in Canada through improved scalability, support for large data files, curation workflows, and integration with Canadian storage and authentication providers.
3. **DuraCloud – Linking Data Repositories to Preservation Storage** – *Co-led by Corey Davis, Council of Prairie and Pacific Research Libraries and Stephen Marks and Kate Davis, University of Toronto*

---

<sup>23</sup> <https://fnigc.ca/ocapr.html>

<sup>24</sup>

<https://static1.squarespace.com/static/58e9b10f9de4bb8d1fb5ebbc/t/5b2973926d2a732f5e620744/1529443237242/2018+IDSov+Global+Progression+-+NAISA+2018+16x9+FINAL.pdf>

- a. Canadian researchers have access to many storage services suitable for the long-term preservation of digital content, including research data. The DuraCloud project will connect several Canadian preservation storage services via this software, which is maintained by the DuraSpace Foundation. As a result, Canadian researchers will be able to seamlessly access different storage services through a single interface.
4. **FAIR Repository for Annotations, Corpora and Schemas (FRACS)** – *Led by André Lapointe, CRIM*
  - a. Artificial intelligence-based applications require access to massive quantities of data. To enable Canada’s academic researchers to scale their AI-based projects such that they are competitive with private sector applications, large volumes of data must be coupled with detailed annotations. Annotated data sets allow models to be effectively trained and validated by machine learning algorithms. The FRACS project will simplify the management of large scale datasets by facilitating the creation, storage, search, manipulation and sharing of their annotations.
5. **Federated Geospatial Data Discovery for Canada** – *Co-led by Eugene Barsky, Evan Thornberry, and Paul Lesack, University of British Columbia Library*
  - a. Traditionally, research data repositories have relied on text-based searching. However, there is increasing demand for geographic components in research, examples of which include migration paths, the distribution of agricultural yields, infrared satellite imagery, the distribution of artifacts in an archaeological site, and the flow routes of water. The goal of this project is to create an extensible, open-source software method to search and discover Canadian geospatial research data using an interface specifically designed for maps, enabling users to discover geospatial resources in a more spatially-intuitive way.
6. **Making Identifiers Necessary to Track Evolving Data (MINTED)** – *Led by Reyna Jenkyns, Ocean Networks Canada (ONC), University of Victoria*
  - a. ONC operates world-leading ocean observatories and dynamic data repository services. While there has been a growing recognition of the benefits and need for data citations made evident by the introduction of the FAIR Principles, existing platforms and tools are currently only able to serve the needs of static or non-frequently updated datasets. The MINTED project will apply best practices for dynamic dataset citation, Digital Object Identifiers (DOIs), and researcher ORCIDs into ONC’s Oceans 2.0 digital infrastructure.
7. **Radium: Management Software for Active Research Data** – *Led by Dr. Kevin Schneider, University of Saskatchewan*
  - a. Research data, which may have value beyond the research for which it was collected, is often distributed across multiple storage devices, tools, and platforms. Simply knowing that a dataset exists, let alone finding it, presents a significant challenge. Radium will provide a project-level metadata index of research data, regardless of where or how it is stored. Radium will improve researchers’ ability to find and cite existing datasets by not only storing the

location of the data, but also the standard and custom metadata records associated with it.

8. **Managing the Research Data Lifecycle using Islandora** – *Co-led by Donald Moses and Rosemary Le Faive, University of Prince Edward Island (UPEI)*
  - a. In collaboration with Simon Fraser University and the Islandora Foundation, UPEI will build research data management capacity and integrations using the latest version of Islandora, also known as CLAW. Islandora is an open-source software framework designed to help organizations collaboratively manage, discover, and share digital assets using a best-practices, standards-based approach. The project will develop integrations with identifier, metadata, authentication, storage, and dissemination systems, supporting the FAIR principles and the research data lifecycle.
9. **Research Portal for Secure Data Discovery, Access and Collaboration** – *Co-led by Dr. Elizabeth Theriault, Ontario Brain Institute and Moyez Dharsee, Indoc Research*
  - a. The Ontario Brain Institute (OBI) and Indoc Research have developed Brain-CODE, an extensible neuroinformatics platform designed to manage the collection, curation, analysis and sharing of different data types across several brain disorders. To address the RDM needs of researchers studying disorders of the brain and other disease areas, this project will develop data portal software that will enable research teams to securely and seamlessly capture, query, and visualize patient data; collaborate and share datasets; and access support and training resources. The project will serve the needs of teams using Brain-CODE as well as those from collaborating institutions and the broader medical research community.

## Evolution of a Canadian NDSF

The model below is based on work by the European Commission<sup>25</sup> to implement a roadmap for the EOSC, and is proposed as a model for the ongoing discussion and development of an NDSF in Canada. The goal with this model is to help frame ongoing discussions at all levels, and to inform the development of a sustainable DRI for DM in Canada. The Canadian DM community has a head-start in developing national DM services, especially with the work of organizations like RDC, CARL Portage, Compute Canada, and the various domain-specific projects and communities of practice. The NDSF conversation, and the adoption of a model like this, allows the full Canadian DM stakeholder community to continue to build on this work.

---

<sup>25</sup> [https://ec.europa.eu/research/openscience/pdf/swd\\_2018\\_83\\_f1\\_staff\\_working\\_paper\\_en.pdf](https://ec.europa.eu/research/openscience/pdf/swd_2018_83_f1_staff_working_paper_en.pdf)

Architecture	Architecture necessary to create a federated system of DM platforms, tools and resources that ensure interoperability and access to all Canadian researchers.
Data	Common standards and tools that support the FAIR principles across disciplines, while ensuring access to rich domain-specific (meta)data.
Services	Generic and discipline-specific services directly supporting researchers and institutions, as well as funder, publisher and institutional policy.
Access & Interface	Human and machines interfaces that make it easy to deposit and access data, including robust support for privacy and security where appropriate.
Rules	Policy and process scaffolding that facilitates participation by all actors in the research ecosystem, supports jurisdictional contexts, and engenders trust.
Governance	Framework that ensures representation by all actors in the development and sustainability of Canada's digital research infrastructure and ecosystem.

The following are offered as examples of conversations/projects that could be used to document in more detail the model described above. These are offered as a way to stimulate the conversation at the NDSF Summit in January.

1. **Architecture:** document architectural details of the 8 featured platforms in the 2019 NDSF Summit, as well as the 9 CANARIE funded projects, with the goal of describing gaps, as well as the steps to achieve a national federation of DM services.
2. **Data:** identify domain repositories that serve as exemplars for specific communities of practice, including a process to submit them to the *CoreTrustSeal*<sup>26</sup> process.
3. **Services:** create a registry of services for researchers to adhere to the components of the Tri-Council DM policy.<sup>27</sup>
4. **Access & Interface:** articulate the value of Canadian data repositories (as an example of a data service) adopting the *SmartAPI*<sup>28</sup>, and work with a few specific use cases to document the effort needed and the impact in making Canada's data services more accessible.
5. **Rules:** draft a policy/process framework for the inclusion of data repositories into an NDSF.
6. **Governance:** draft an approach to governance for an NDSF, that ensures input from the complete stakeholder community.

<sup>26</sup> <https://www.coretrustseal.org/>

<sup>27</sup> [https://www.science.gc.ca/eic/site/063.nsf/eng/h\\_97610.html](https://www.science.gc.ca/eic/site/063.nsf/eng/h_97610.html)

<sup>28</sup> <https://smart-api.info/>

## Next Steps

Canada is by no means starting from scratch when it comes to building an NDSF, as a great deal of work has been done to bring a strong suite of research platforms and services to Canadian researchers. The goal of a go-forward strategy should be to build on this existing effort to fill gaps and ensure the ongoing development and long-term sustainability of these resources.

The January 2019 NDSF Summit is the next step in this conversation, with the proposed outcomes of:

1. developing a more detailed description of the current Canadian NDSF landscape, including intersections and interoperability of projects;
2. moving ahead with six pilot “projects” to intersect with the 6-part NDSF model;
3. developing recommendations for moving forward and reporting at the 3rd NDSF Summit.

The 2019 Summit will feature presentations from 8 leading/emerging Canadian DM platforms and tools, viewed through the lenses of the FAIR Principles. These presentations will stimulate the conversation regarding the Canadian DM ecosystem, and how to best move forward to meet the needs of all Canadian researchers. The 8 presentations and associated projects are listed below.

- 1. Findability**
  - a. FRDR: Federated Research Data Repository (Portage/CC)
  - b. Epigenomic Data Discovery with the IHEC Data Portal (McGill University)
- 2. Accessibility**
  - a. Secondary Analysis to Generate Evidence, SAGE (PolicyWise)
  - b. Canadian Research Data Centre Network CFI Project (CRDCN)
- 3. Interoperability**
  - a. Canadian Integrated Ocean Observation System (CIOOS)
  - b. Canadian Astronomy Data Centre (CANFAR CADC)
- 4. Reusability**
  - a. DataStream (Gordon Foundation)
  - b. SPOR (Strategy for Patient-Oriented Research) National Data Platform (CIHR)

The Summit will include breakout sessions designed to provide recommendations for next steps, which will be further discussed by the stakeholder community after the Summit to help determine how to move forward. The discussion emerging from these meetings will be summarized in a report and made available to the broad stakeholder community. It is anticipated that these efforts will intersect with the direction articulated by ISED in the coming months, and help the DM community “hit the ground running” when this clarity and direction is provided.

# Resources

Additional resources.

- [Implementation Roadmap for the European Open Science Cloud](#)
- [EOSC Declaration](#)
- [Mapping the Open Science Ecosystem](#)
- [NLM Strategic Report](#)
- [NIH Strategic Plan](#)

# Appendix A: The Vienna Declaration on the European Open Science Cloud

Vienna, 23 November 2018

## **We, Ministers, delegates and other participants attending the launch event of the European Open Science Cloud (EOSC):**

1. **Recall** the challenges of data driven research in pursuing excellent science as stated in the "EOSC Declaration" signed in Brussels on 10 July 2017.
2. **Reaffirm** the potential of the European Open Science Cloud to transform the research landscape in Europe. Confirm that the vision of the European Open Science Cloud is that of a research data commons, inclusive of all disciplines and Member States, sustainable in the long-term.
3. **Recognise** that the implementation of the European Open Science Cloud is a process, not a project, by its nature iterative and based on constant learning and mutual alignment. Highlight the need for continuous dialogue to build trust and consensus among scientists, researchers, funders, users and service providers.
4. **Highlight** that Europe is well placed to take a global leadership position in the development and application of cloud services for Science. Reaffirm that the European Open Science Cloud will be both European and open to the world, reaching out over time to relevant global research partners.
5. **Recall** that the Council - in its conclusions of 29 May 2018 - welcomed the implementation roadmap and the federated model for the European Open Science Cloud. It invited the Commission and all Member States to set up a common governance framework that ensures participation of stakeholders from the research community based on principles of transparency, openness and inclusiveness and an effective involvement of all Member States.
6. **Note** that the 2018 EOSC Summit (held on 11 June 2018) called for acceleration towards making the European Open Science Cloud a reality, hinting at the need to further strengthen the ongoing dialogue across institutions and with stakeholders, for a new governance framework to be launched in Vienna, on 23 November 2018.

## **We therefore:**

7. **Resolve** to harness the many ongoing and planned activities at EU and Member States level to cooperate in establishing an inclusive partnership with a view to developing the European Open Science Cloud as a federated infrastructure that can enhance value-based, open, trusted, user-centric digital services across borders within the Digital Single Market (DSM).
8. **Invite** all Member States, as well as public and private stakeholders in Europe, to support actively this joint effort and the new European Open Science Cloud governance structure for a successful implementation of the initiative.
9. **Call** for the European Open Science Cloud to provide all researchers in Europe with seamless access to an open-by-default, efficient and cross-disciplinary environment for

storing, accessing, reusing and processing research data supported by FAIR data principles.

10. **Commit** to support service provision for the European Open Science Cloud by helping connecting relevant national and disciplinary nodes to the pan-European level.
11. **Reaffirm** the potential of the European Open Science Cloud to enable first-class data-driven science and to stimulate new business models benefiting our society and the economy. **Recognise** that such services will create opportunities for both public and private sectors, notably by intensifying reuse of public sector information while preserving data integrity, and ensuring access, transparency within and across borders.

**We therefore declare** to work together towards realising the potential of the European Open Science Cloud for the benefit of citizens, society and the economy.

*The Vienna EOSC Declaration was conceived by Paolo Budroni, University of Vienna and Stefan Hanslik, Austrian Federal Ministry for Education in close cooperation with the European Commission. Vienna, 2018*

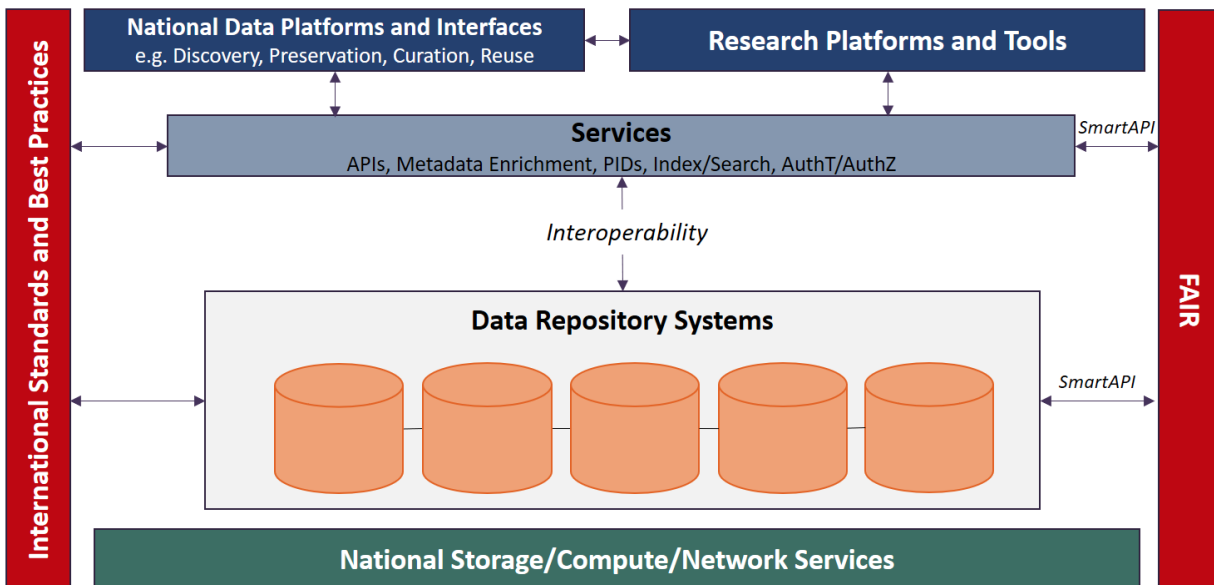
Contact

Launch of the European Open Science Cloud (EOSC)  
University of Vienna  
Library, main reading room  
Universitätsring 1  
A-1010 Vienna  
T: +43-1-4277-17677  
[congress@univie.ac.at](mailto:congress@univie.ac.at)

Original Declaration is [available to download](#).



## Appendix B: NDSF Diagram



This high-level view of a NDSF was included in the FAQ for the 1st CANARIE RDM Call, and is intended to show the various layers of infrastructure and services that are critical to an effective national framework. A more detailed version of this that reflects specific projects that function at a national level is being developed using the 9 CANARIE-funded projects as the context.