National Data Services Framework Summit (NDSF): 2020 Summary and Outputs

Robyn Nicholson| Intern, Research Data Canada @robyncecelia
Mark Leggott | Executive Director, Research Data Canada @rdc_drc

Agenda

> Background of the NDSF and Summit

> NDSF Summit 2020 Overview

> NDSF 2020 Summary and Recommendations Document

> Kanata Declaration 2020

> Looking Ahead to NDSF Summit 2021

> Questions and Answers

The Outputs

English:

> National Data Services Framework (NDSF) Summit 2020: Summary and Recommendations: https://doi.org/10.5281/zenodo.4002141

> Kanata Declaration 2020: https://doi.org/10.5281/zenodo.4002076

French:

> Sommet de 2020 sur l'encadrement des services de données nationaux (ESDN) : Sommaire et recommandations: https://doi.org/10.5281/zenodo.4006502

> Déclaration de Kanata 2020: https://doi.org/10.5281/zenodo.4004831

RDC ✦ DRC
Research Data Canada · Données de recherche Canada

Slide Decks

English: bit.ly/RDC_NDSF_EN

**French**: bit.ly/RDC_NDSF_FR

Simultaneous Interpretation in French

> This session features both English and French language audio channels.

> *Cette session propose des canaux audio en anglais et en français.*

> Select the Interpretation option from your menu to select preferred language.

> *Sélectionnez l'option Interprétation dans votre menu pour sélectionner la langue préférée.*

> A recording of this session, along with these slides, will also be made available in French.

> *Un enregistrement de cette session, accompagné de ces diapositives, sera également disponible en français.*
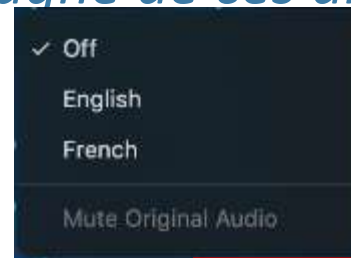
Questions & Answers

> Please use the Q&A option to ask questions of the presenters. Questions will be monitored throughout the session and addressed at the end.

> The Q&A option can be found at the bottom of your Zoom screen:
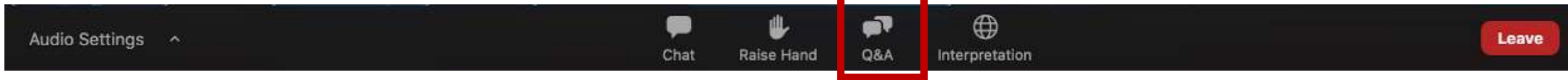
> Please note that this **event is being recorded**, including questions and answers.

Background

NDS/NDSF

> National Data Service
  - A service that provides one or more data-related functions to applicable stakeholders and disciplines in a specific national context.

> National Data Services Framework
  - a conversation with all stakeholders at all levels;
  - agreement on best-practices, standards and protocols;
  - a suite of interoperable services and resources.
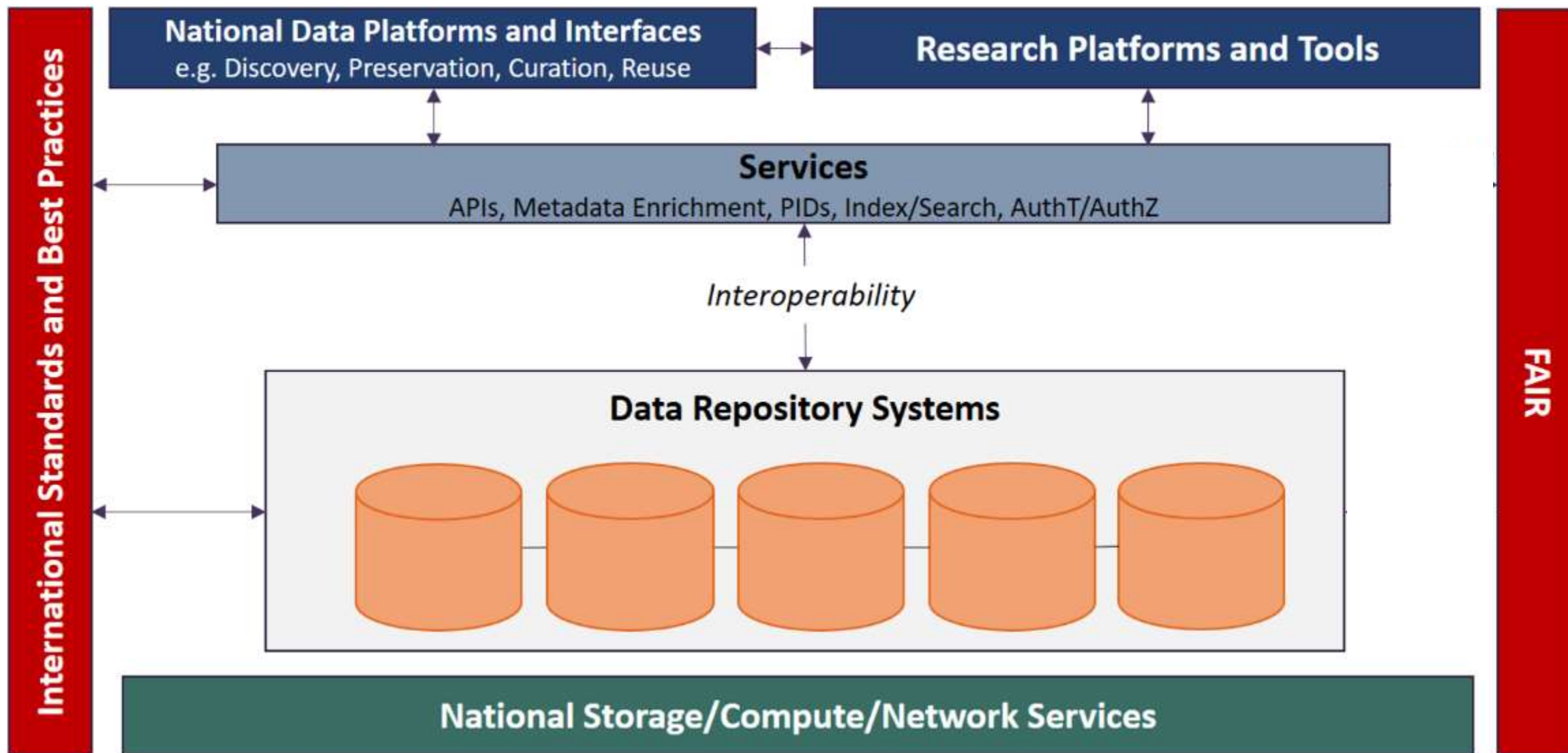
Previous NDSF Summits

> 1st Summit held in conjunction with the RDA Plenary on Montreal
  - September 22, 2017
  - Leveraged international stakeholders for sessions on NDSs in general
  - Introduced concept of NDSF, and highlighted national and international efforts

> 2nd Summit held in Kanata, Ontario
  - January 24-25, 2019
  - Highlighted national efforts to build NDSs
  - 1st Kanata Declaration emerged from the discussion

> 3rd Summit also in Kanata
  - February 5-6, 2020
  - Discussion of existing, emerging, needed NDSs, and defining parameters for NDSs

RDC ✦ DRC

NDSF Summit 2020 Overview

Who Were We?
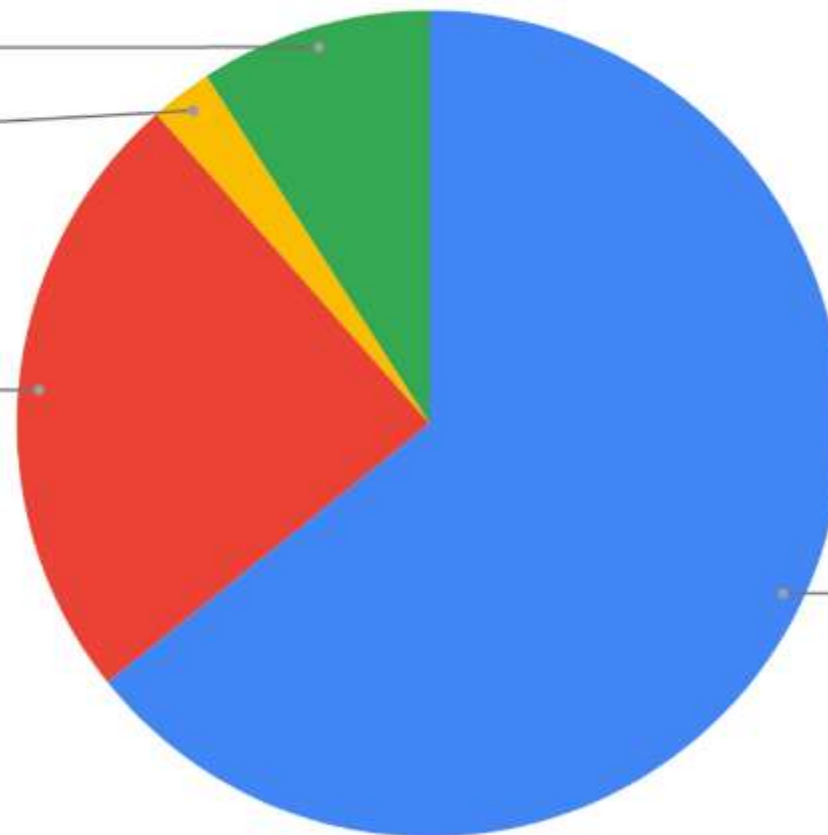

Type of organization

- **Federal Government** 9.1%
- **Private** 2.4%
- **Non-Profit** 24.2%
- **Academia** 64.2%

> Over 160 Attendees from over 80 different institutions!

RDC ❧ DRC

Who Did We Represent?

Attendee's Association

Chart showing Organizational Units: IT 30, Funder 16, Data Org 25, Service Org 18, Researcher 19, Library 40, Research Office 17
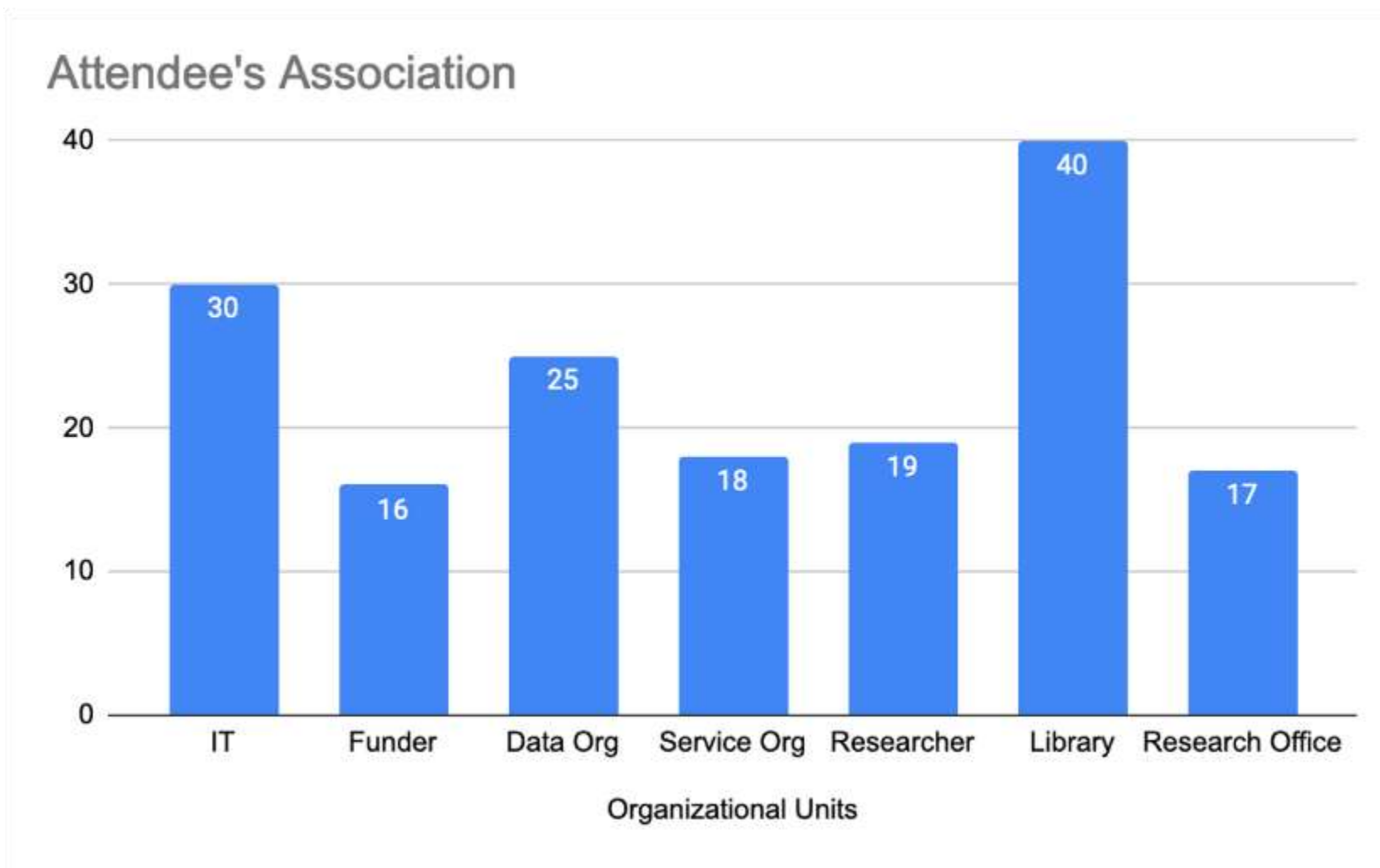
NDSF 2020 Goals and Outcomes

> Consensus on what minimal requirements define a NDS;

> Consensus on what works best at the institutional, provincial, regional, national, and international levels;

> Consensus on existing Canadian NDSs, both machine and human services, that provide either models for, or operational examples of NDSs;

> Consensus on key gaps not articulated in either the Research Data Management Roadmap or Kanata Declaration;

> Consensus on priorities going forward.

NDSF Summit 2020 Summary and Recommendations

NDSF Summit 2020 Breakout Groups Input

> Attendees participated in three breakout sessions and were asked to consider the following topics:
>   1. requirements for a National Data Service (NDS);
>   2. examples of NDSs (existing, emerging, and needed);
>   3. priorities for research data management (RDM) moving forward (the next 6 and 12 months, 2 and 5 years).

> Feedback from these breakout groups was transcribed, analysed, and rolled up into a summary which formed the basis of the NDSF Summit 2020 Summary and Recommendations document.

Breakout Chart Transcription

Overarching Themes

> Overarching themes emerged across all three breakout sessions, including:

> > Communication and training
> > Alignment and integration of existing organizations/services
> > Centralization and federation of Canadian of RDM components
> > Balance between domain-specific and domain-agnostic services
> > Long-term preservation of data
> > Management of sensitive data

RDC ❖ DRC

Discussion Framework

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.

> Feedback was then categorized based on the European Open Science Cloud (EOSC) 6-part model that defines 6 "lines of action" for facilitation, adapted here for the Canadian context.

RDC ♦ DRC

Discussion Framework

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.

Communication

Training

RDC ☀ DRC
Research Data Canada · Données de recherche Canada

Category Highlights

Architecture
- **User-centred design** with accomodation for all sectors.
- **Centralized federated architecture** with local infrastructure, accessible through API defined by common standards.

Data
- Universal terminology with **multilingual support**.
- **Common data models** integrated with domain-specific (meta)data standards to serve data across disciplines.

Services
- **National training as scaffolding** for local skills development, with implementation/technical advisory teams.
- Collaboration of RDM stakeholders to provide **enhanced and open access to coordinated services**.

Access & Interface
- **Incentivize best practices** for researchers to standardize deposit processes.
- API or linked APIs with flexibility and **data discovery, inventory, and exploration tools**, and identity management.

Rules
- Canadian **common baseline** for standards and policies.
- **Transparency** in policy and decision-making for key service parameters.

Governance
- Active engagement, **inclusion, and representation** of all sectors and domains in RDM governance.
- Nationally funded DSF with **international interoperability** and ongoing regional/local consultation.

RDC �▲ DRC

MindMap Visualization
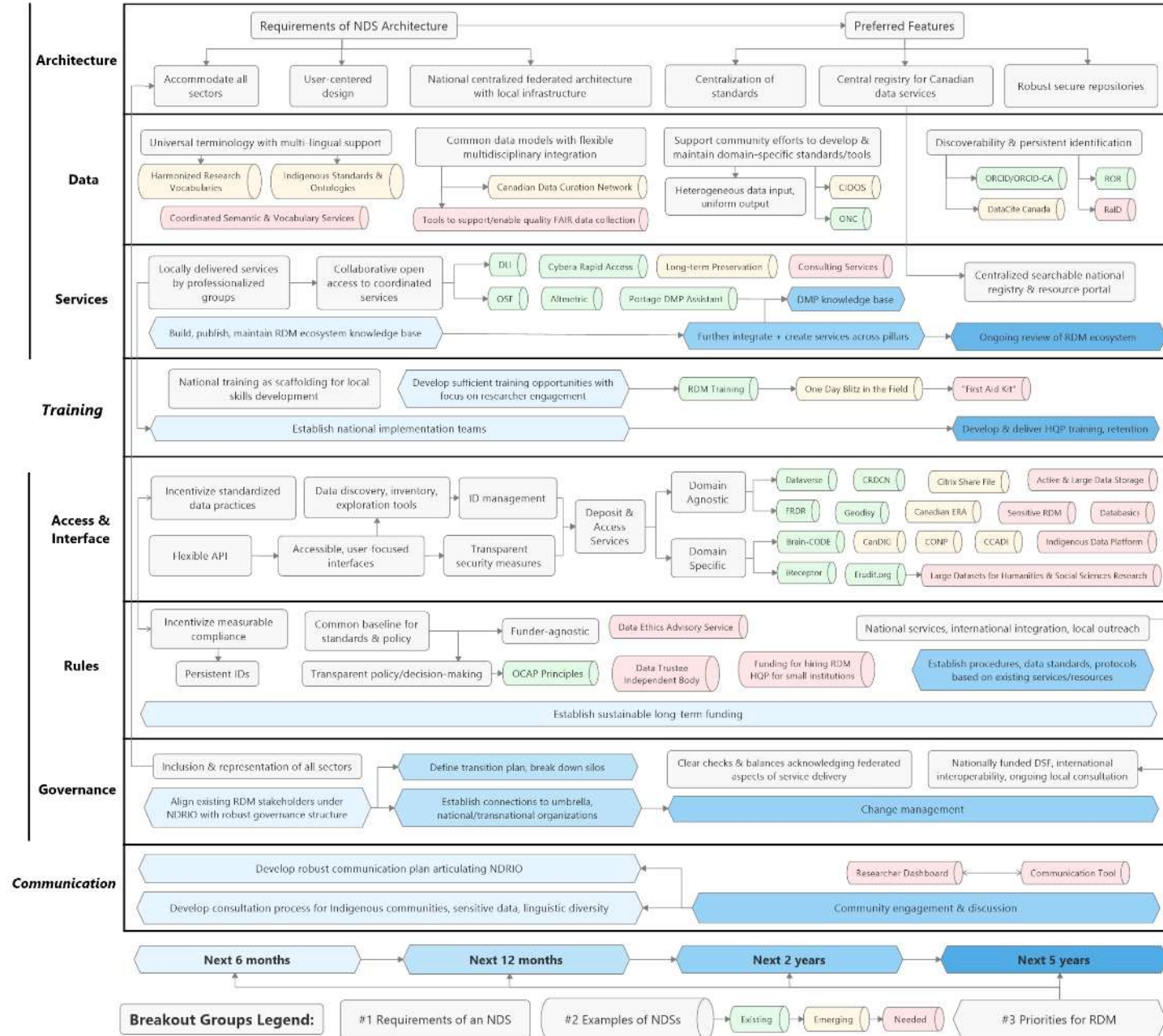
MindMap Visualization
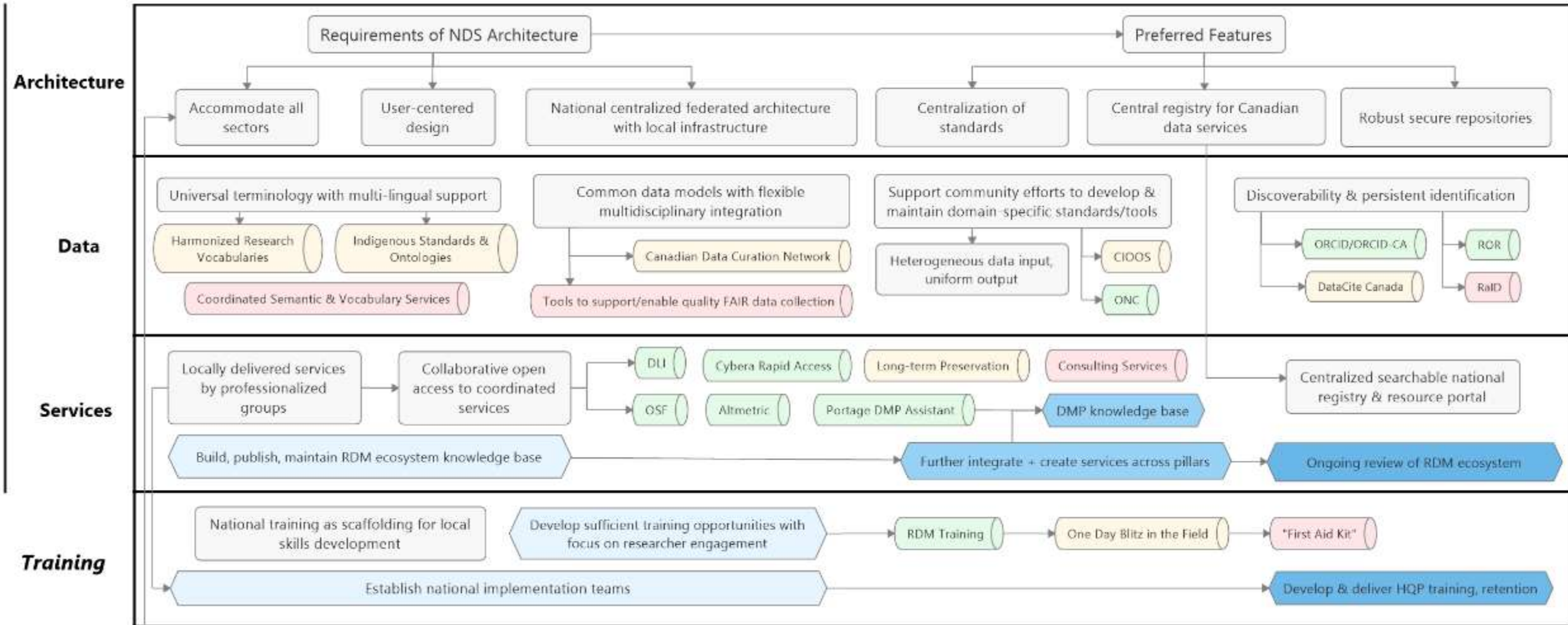
Recommendations

> A total of 40 recommendations were extracted from participant feedback and aimed at broad stakeholder groups, including:

> > Researchers
> > Research Funders
> > Universities and Research Centres
> > Science-based Government Departments and Agencies (SBDAs)
> > Repositories and Publishers
> > RDC, Portage, and NDRIO

Recommendation Highlights for All Stakeholders

1. Continually **communicate** the developments of the RDM ecosystem to researchers and the general public.

2. Develop and deliver a national **training** program that serves as scaffolding for local skills development.

3. Accommodate all sectors and domains in terms of national architecture and **diverse inclusive representation** in governance.

4. Through consultation, develop more **services focused on Indigenous data**, including platforms, repositories, standards, and vocabularies.

5. Use and develop terminology and controlled vocabularies with **multilingual support** to improve access and interoperability.

6. Develop **long-term data preservation** options, potentially in the form of a national data preservation infrastructure.

7. Establish **funding for hiring RDM HQP**, particularly for smaller institutions.

RDC 🍁 DRC

Kanata Declaration 2020

Kanata Declaration 2019

> Kanata Declaration named after community 1st meeting held in
> Declaration framed statements using keywords: Recognize, Highlight, Resolve, Commit, Call
> Extracted from dialog and breakouts at NDSF Summit 2
> Goal to create a set of Principles to:

- articulate primary interests of researchers in the context of data management
- frame ongoing discussions to build NDSs to support researchers
- facilitate efforts to develop "rules of participation" for NDSs
- report back to the community on efforts/activities in the context of each statement

RDC 🍁 DRC

Kanata Declaration 2020

> 2nd version of the Declaration reflected feedback from
>   - Survey Monkey form
>   - Public Feedback period for 2019 Declaration
>   - Discussions at the NDSF Summit 2020
>   - Public Feedback period for 2020 Declaration Draft

> 24 statements provide a fulsome, principles-based framework that reflects the voice of the full RDC stakeholder community
>   - researchers
>   - funders
>   - policy makers
>   - infrastructure providers
>   - librarians/data managers/curators/experts

Collaborative and Community-Driven

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 standardsbest practices~~metadata standards such as schema.org, ~~Common Archive Observation Model (CAOM)~~, ISO 19115, and the DataCite Metadata Schema, or on an appropriate linked data model with support for domain-specific metadata.

Resolve to ensure that ~~First Nations, Métis, Inuit and broader global Indigenous communities Indigenous communities~~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, including~~especially~~ via the:

1. ~~the~~consideration of Indigenous data and Traditional Knowledge (TK)~~;~~,
2. ~~-~~use of Indigenous-specific approaches ~~((~~such as the OCAP and CARE Principles~~)~~;
3. use of and mechanisms ~~(~~for clarifying indigernous-specific language regarding use and access, such as the TK ~~Labels~~ and Biocultural (BC) Labels);
4. development of specific meta-data fields that recognize Indigenous interests, supporting better provenance practices for Indigenous data, and encourag~~e~~ing use by researchers of new notification tools (e.g. TK and Biocultural (BC) Notices) participatory meetings and gatherings.

Highlights of 2020 Declaration

> 1st Principle provides the tone for the rest of the document.

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 community, and creates an environment where all research can be questioned and advanced.*

Highlights in the Context of the 6-part EOSC Model

1. Architecture
   a. *Highlight: 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.*

2. 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.*

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

Principles in the Context of the 6-part EOSC Model

4. Access and Interface
   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.*

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

6. Governance
   a. *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.*

Indigenous Data Sovereignty/Considerations

> **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:

- use of Indigenous-specific approaches such as the OCAP and CARE Principles;
- 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.

Privacy and Ownership Considerations

> *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).*

> *Commit to ensuring that the control and ownership of their 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.*

Responsibility to Report back to the Community

> ***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.*

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

> **Reporting Matrix coming later in 2020**

Ultimate Goal of the Kanata Declaration

> To provide a high-level, community-crafted, principles-based approach to ensuring researchers receive the best support possible for data sharing and stewardship…

> and to provide a foundation to go back to when you stuck in the weeds!

RDC ✦ DRC
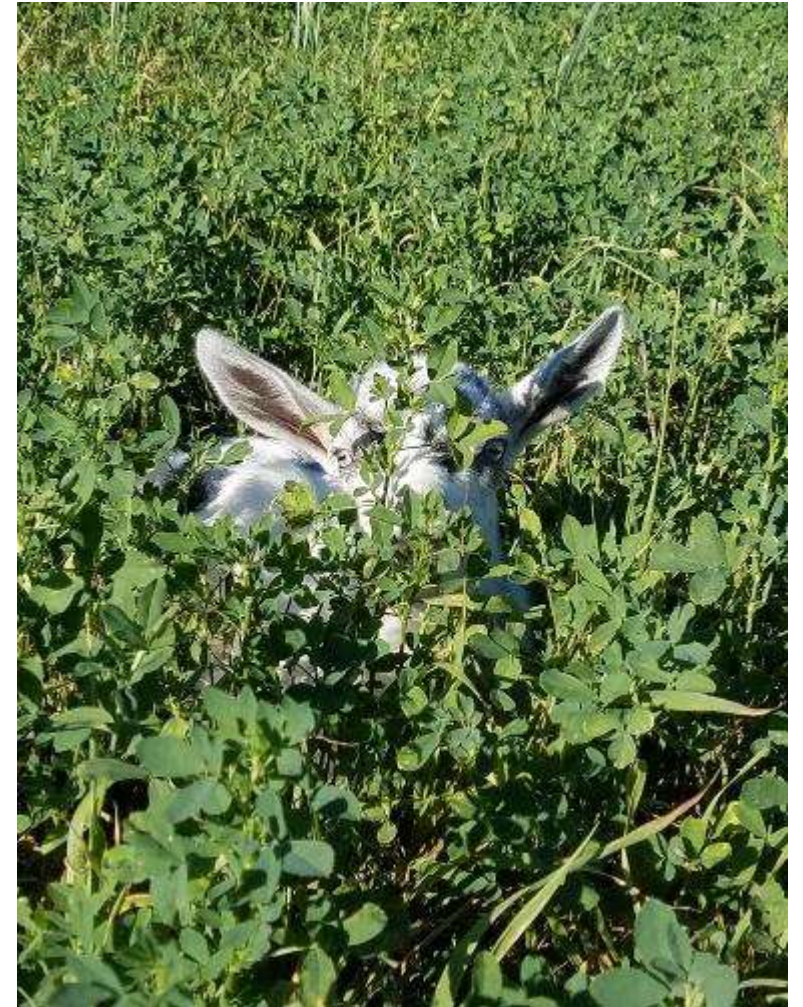
NDSF Summit 2021

NDSF Summit 2021

> Planning for February 2021
  - Virtual Summit with breakouts to focus on details of the NDSs
  - Suggestions for approaches/themes are welcome
  - Breakout themes will derive from the Kanata Declaration, and develop actionable objectives
  - Informing NDRIO in the transition period will also be a focus

> Lead-up to the 2021 Summit
  - Presentation of a reporting matrix based on 2020 Kanata
  - Background documents, including NDRIO Data Management & other Position Papers

RDC ✦ DRC

Thank you! Questions?

Mark Leggott, Executive Director @rdc_drc - mark.leggott@rdc-drc.ca
Robyn Nicholson, Intern @robyncecelia - robyn.nicholson@rdc-drc.ca