

The Road to Open Science is Paved with Canadian Bedrock

Mark Leggott | BCNET | April 24, 2018

Let's connect: mark.leggott@rdc-drc.ca | [@rdc_drc](https://twitter.com/rdc_drc)



Research Data Canada – Données de Recherche Canada

rdc-drc.ca | [@rdc_drc](https://twitter.com/rdc_drc)



Value Proposition

- > Whale Fall = Canvas for Evolution
 - Is non-discriminatory re who/what can benefit
- > Data = Canvas for Innovation & Discovery
 - Currently not available nor non-discriminatory



Biodiversity



Dataversity



7 Reasons for Not Sharing Data

1. Will be disadvantaged in academic rat race
2. A researcher will “scoop” me using my data
3. Data may not be replicable
4. Another will come up with different results
5. Data is poor/in poor shape
6. The data is confidential
7. Technical challenges to sharing data

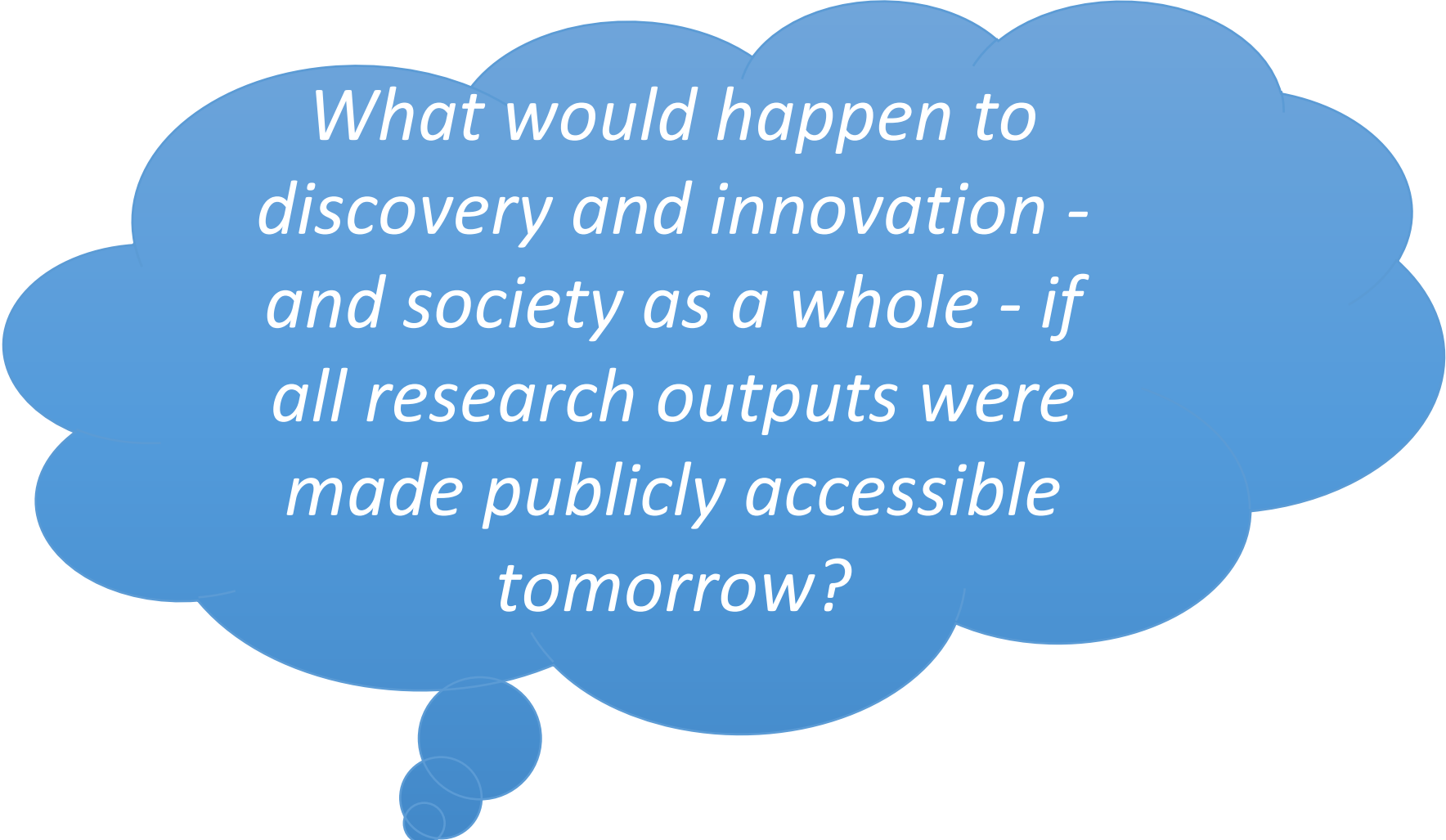
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4909097/>



FAIR

- > Findable
 - PID and rich metadata; Indexed and linked to data
- > Accessible
 - Retrievable; Best Practices; AuthT/AuthZ if needed
- > Interoperable
 - Standard (meta)data, data formats & vocabularies
- > Reusable
 - Good (meta)data; clear license; domain-relevant





*What would happen to
discovery and innovation -
and society as a whole - if
all research outputs were
made publicly accessible
tomorrow?*

“If 100 people look at the data that I work with, some of them will be interested in different things than I’m interested in. It’s not that I don’t think I’m smart, but I have my own particular point of view. Other people looking at the same data would ask different questions than I would and that might lead to something new.”

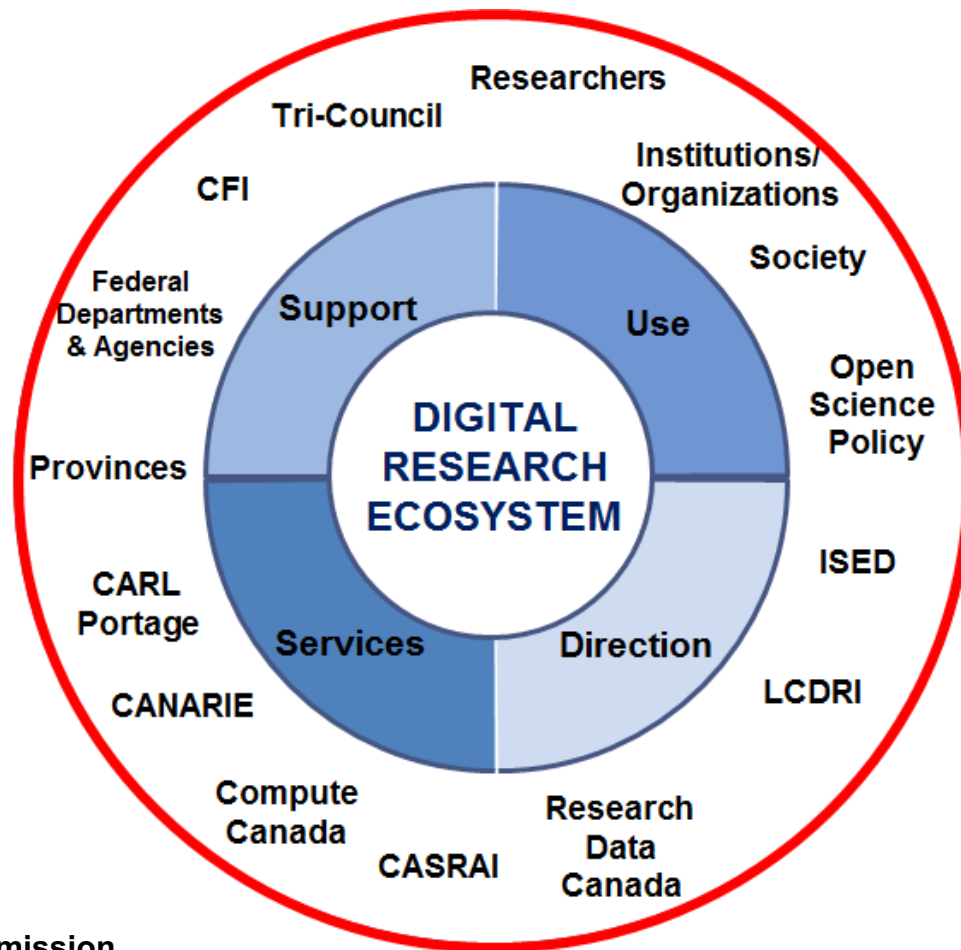
Guy Rouleau, Director, Montreal Neurological Institute





> **Canadian Landscape: An Update**





Adapted from CFI DRI submission



Canada's "Open" Focus

- > Federal government and provinces increasingly committed to "open", including Open Science
 - Canada co-chairing Open Government Partnership
 - G7 Science Ministers and Open Science/DRI
 - Chief Science Advisor, Open Science lead
- > 2018 Budget commitment to research, including \$572 million for DRI
 - \$4 billion funding for research in various contexts



Research Data Canada

- > RDC works with stakeholders to ensure research data is available to support innovation that benefits all.
- > **Vision:** Canadians have access to national and international collections of publicly-funded research data.
- > **Mission:** Work with stakeholders to ensure research data is re-usable in support of innovation that benefits all Canadians.



CANARIE and NREN Partners

- > CANARIE + 12 provincial and territorial partners = Canada's National Research & Education Network
- > Delivers digital infrastructure that supports world-class research and innovation
- > 1st RDM Funding Program launching in 2018
 - May-June: Submission timeline
 - October: Funding
 - March 2020: End of Program



International RDM Infrastructure

- > World Data System (ICSU) International Technology Office launched in March 2018
- > U of Victoria, collaboration with ONC, CADC, CCIN/PDC
 - Operate & manage contributions of WDS to the Global Research Data Infrastructure (GRDI)
 - Coordinate development & integration of GRDI components with other international frameworks
 - Host a Trusted Brokering Service for global community
 - Coordinate WDS contributions to technical WGs



Tanenbaum Open Science Institute

- > Seeded by \$20 million donation from Lawrence and Judith Tanenbaum Family Foundation
- > Part of MNI and Faculty of Medicine
- > OS as a way to achieve “accelerated discovery, new treatments, economic growth and healthier population”.
 - Expand impact of MNI data via global data sharing
 - Develop tools and DRI to support data/material sharing
 - Measure impact of open science
 - Encourage scientific community to embrace OS



Tri-Council DM Policy

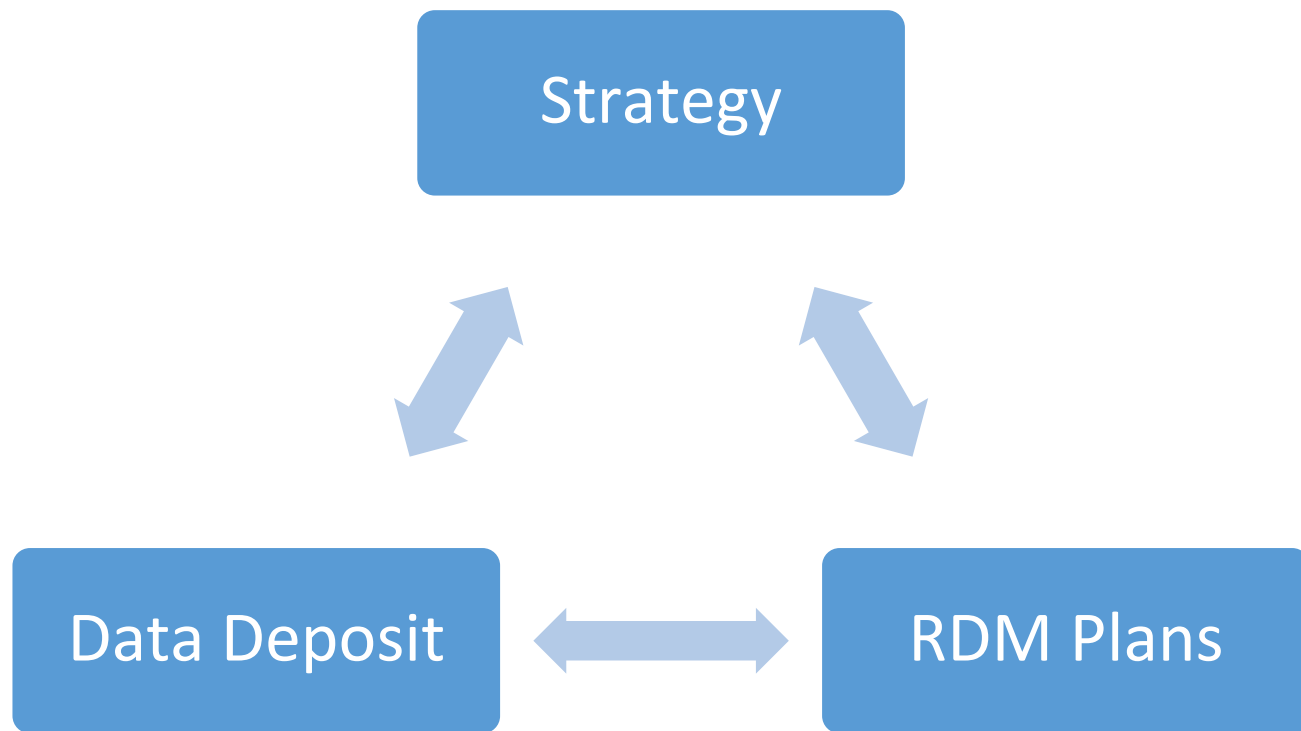
- > 2013: *Capitalizing on Big Data: Toward a Policy Framework for Advancing Digital Scholarship in Canada*
- > 2014-2016: Discussion leading to development of *Statement of Principles on Digital Data Management* in 2016
- > 2017-2018
 - Announcement of draft policy at CARA (May/17)
 - TC3 Policy Roadshow, west to east (Jan-Feb/18)
 - Internal “tweaks” to Policy draft given stakeholder feedback (Mar-Apr/18); launch of draft coming



Three Pillars of the TC3 Policy

- > Institutional Strategy: Develop a local strategy/policy for RDM
 - **What Works for Institutions**
- > Data Management Plans: Researchers create plans for (some) TC3-funded projects
 - **Internal Conversation re: Needs**
- > Data Deposit: Deposit data (and associated code/resources) in a repository/platform
 - **Start With Data [Availability](#)**







SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

The Road to Open Science is Paved with Canadian Bedrock

Shahira Khair | BCNET | April 24, 2018

shahira.khair@carl-abrc.ca
@portageCARLABRC

What is the Portage Network?

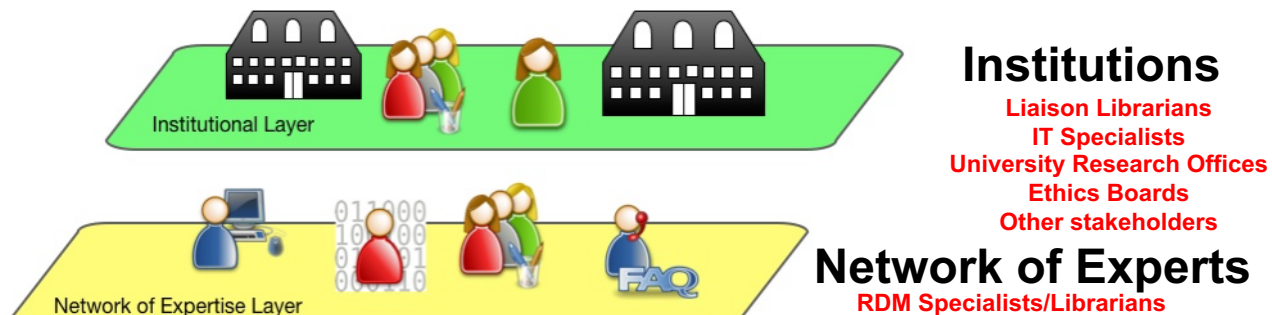
“Portage is a national, library-based research data management network that coalesces expertise and initiatives to develop Canadian capacity”



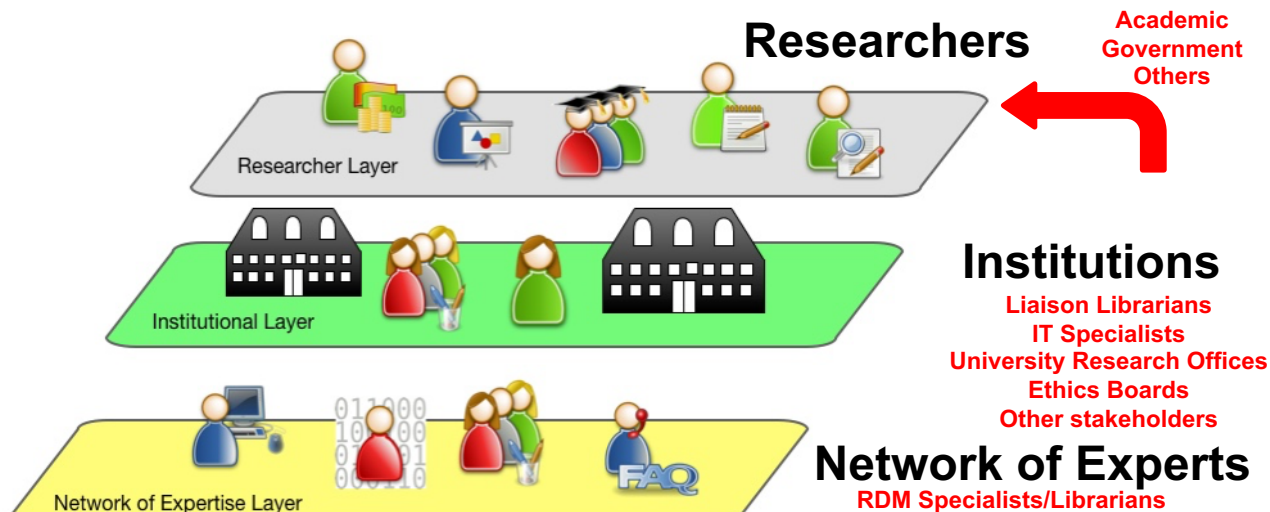
Goals of the Portage Network

- Build a **community of practice** for research data management
- Engage and advocate for research data management with **stakeholder communities**
- Facilitate and provide leadership in the **development of RDM infrastructure**

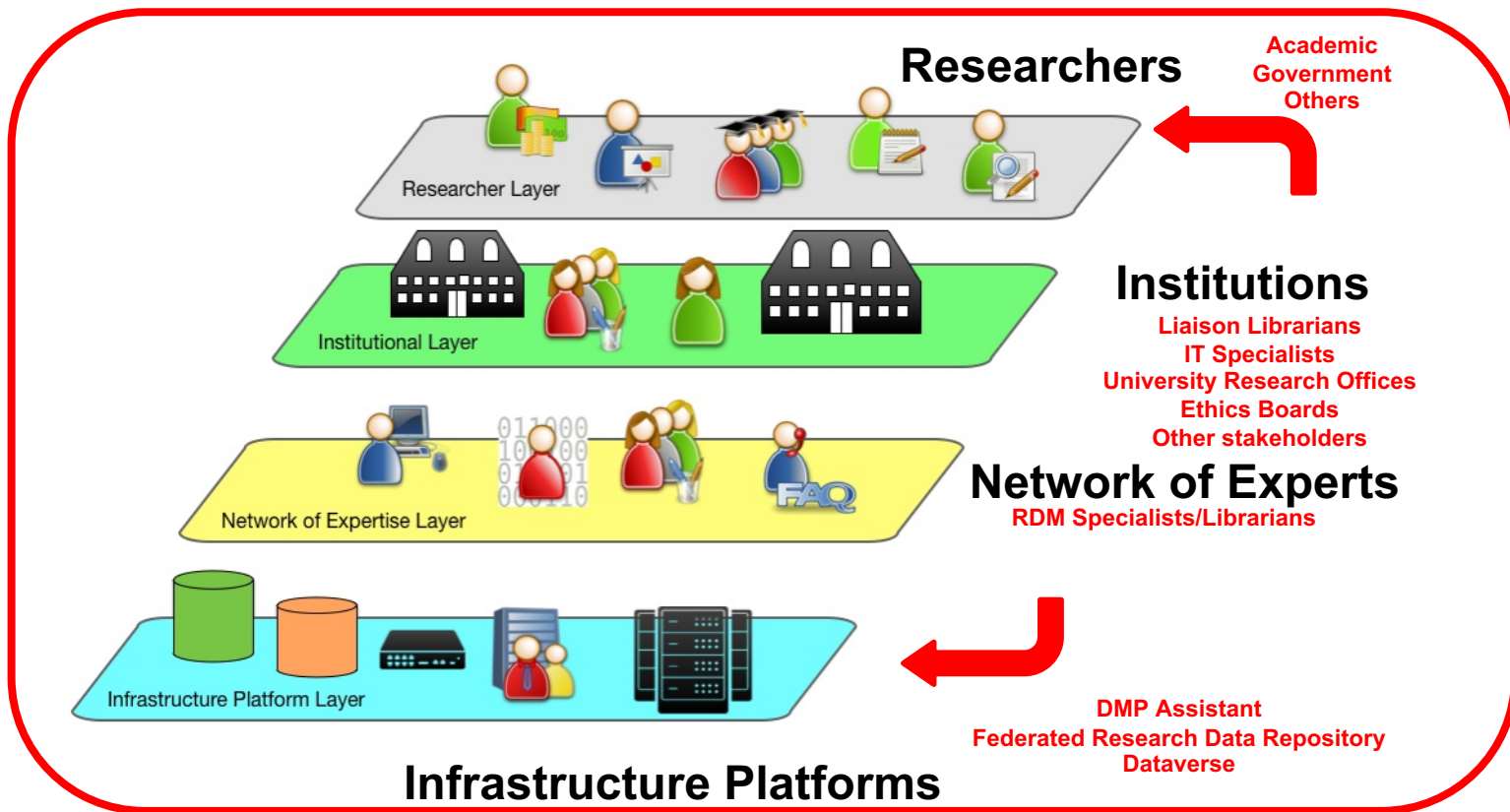
Building the Portage Network



Building the Portage Network



Building the Portage Network



Drivers in the RDM Landscape

Institutional
Strategy



DMPs



Deposit



But the work of Portage is also informed by and aligned with trends in journal publishing, open science, international RDM initiatives and recognized best practices

Goals of the Portage Network

- Build a **community of practice** for research data management

Portage Network of Experts



100+

Volunteers



40+

Organizations

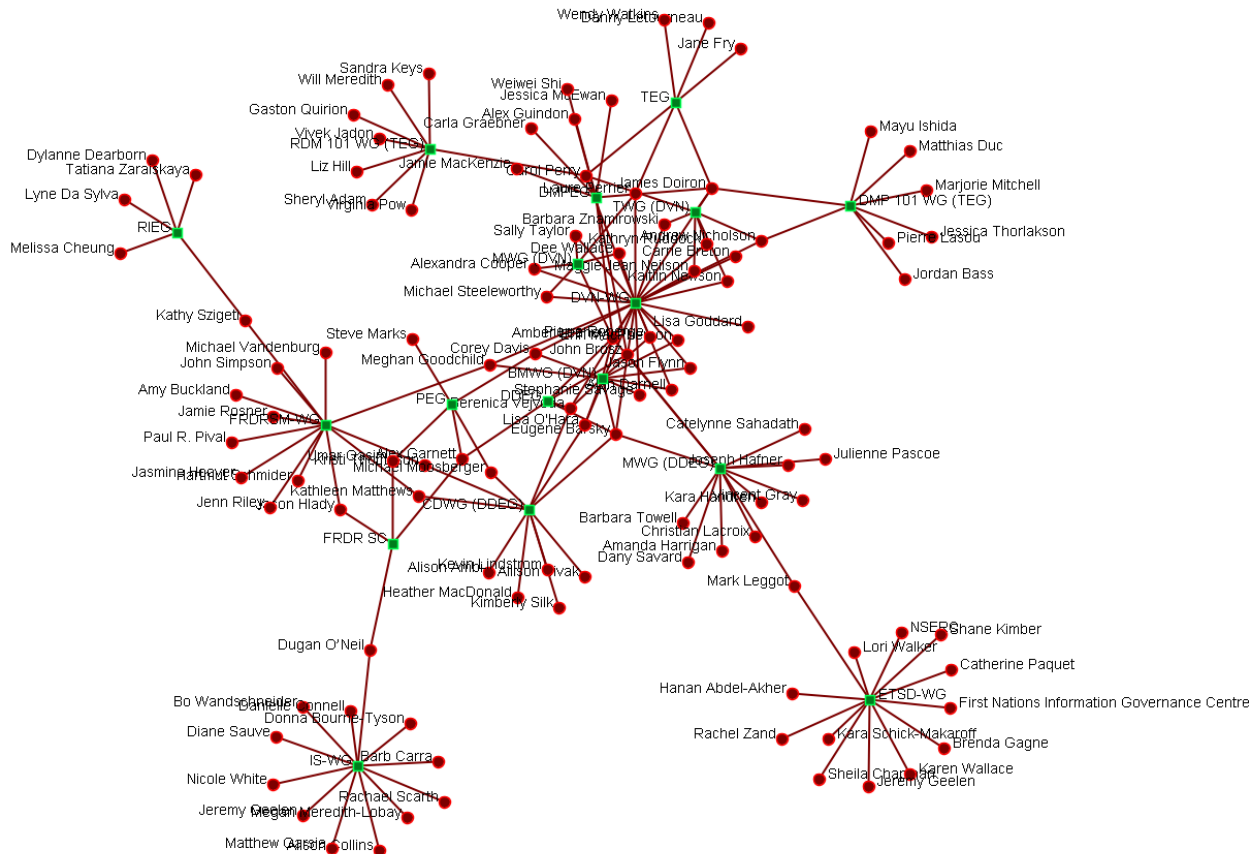
Expert Groups

- RDM Planning
- Curation
- Data Discovery
- Preservation
- Training
- Research Intelligence

Working Groups

- Repository Platforms
- Institutional Strategies
- RDM Practices for Sensitive Data

A True Network of Experts!



Guidance and Resources

Research Data Discovery and the Scholarly Ecosystem in Canada A White Paper

Prepared by the Port
Canadian Association
Eugene Berris (Concordia)
John Bross (University of
Alberta) Lesley (York)

JULY 2016

Research Data Management Training Landscape in Canada A White Paper

Prepared by the Portage Training Expert Group
on behalf of the Canadian Association of Research Libraries (CARL)

Jane Fry (Carleton University)
James Down (University of Alberta)
Dany Liouane (Université de Moncton)
Carol Perry (University of Guelph)
Wendy Williams (Carleton University, retired)

www.carl-abrc.ca

JANUARY 2017

Portage Network
Canadian Association of Research Libraries
portage@carl-abrc.ca

www.carl-abrc.ca

Portage Data Discovery Expert Group - Phase One Report

Work Collections Development Working Group
on behalf of the Canadian Association of Research

1) Chair
(New Brunswick)
Columbia)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

(Y)
(Y)
(Y)

RESEARCH DATA MANAGEMENT PRIMER

Research Data Management (RDM)

- RDM refers to the processes applied throughout the lifecycle of a research project to guide the collection, documentation, storage, sharing, and preservation of research data.
- RDM practices are integral to conducting responsible research and can help researchers save resources by ensuring their data is complete, understandable, and secure.
- RDM practices also follow institutional and funding agency guidelines that protect their investments.
- The broader research community can derive maximum value from research data that can be accessed, shared, reused and repurposed.

The Research Lifecycle



Defining Research Data

- Primary sources supporting research, scholarly research, or other research.
- Can be used as evidence to validate findings.
- May take the form of experimental data, survey data, public sector data, monitoring data, etc.
- All other digital and non-digital content that is research data.

*Research data: (OCLC) In CARL's Dictionary, RDM refers to the lifecycle of research data.

Portage Network

Canadian Association of Research Libraries

portage@carl-abrc.ca

www.carl-abrc.ca

Portage Network

Canadian Association of Research Libraries

portage@carl-abrc.ca

www.carl-abrc.ca

Portage Network

Canadian Association of Research Libraries

portage@carl-abrc.ca

www.carl-abrc.ca

Portage Network

Canadian Association of Research Libraries

portage@carl-abrc.ca

www.carl-abrc.ca

Portage Network

Canadian Association of Research Libraries

portage@carl-abrc.ca

www.carl-abrc.ca

Portage Network

Canadian Association of Research Libraries

portage@carl-abrc.ca

www.carl-abrc.ca

"Good Enough" Research Data Management (a brief guide for busy people)

This brief guide presents a set of good data management practices that researchers can adopt, regardless of their data management skills and levels of expertise.



Online Modules *Coming Soon...*

- RDM 101
- Data Management Planning

portagenetwork.ca

Institutional RDM Strategy Template

Strategy Components

1. Raising awareness
2. Assessing institutional readiness
3. Formalizing RDM practices
4. Defining a roadmap

Institutional
Strategy



DMPs



Deposit



Goals of the Portage Network

- Build a **community of practice** for research data management
- Engage and advocate for research data management with **stakeholder communities**

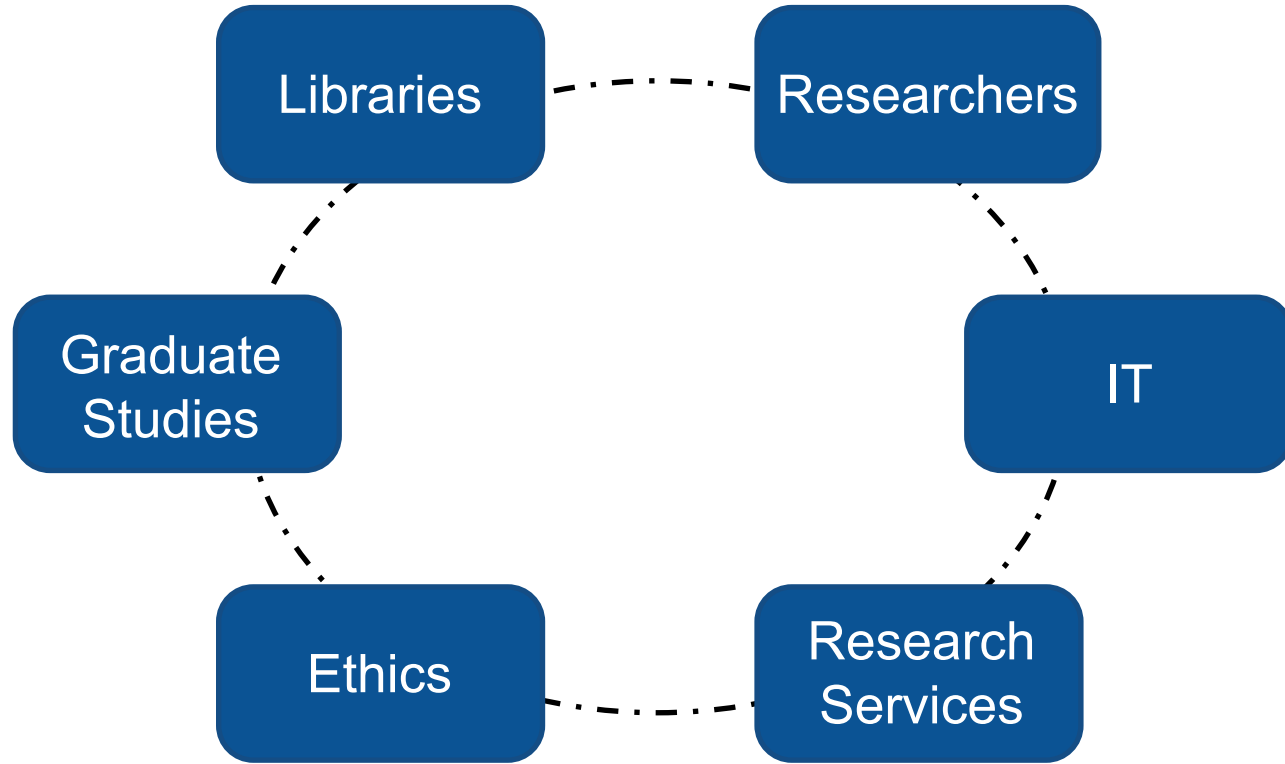
Levels of Stakeholder Engagement

LOCAL

**REGIONAL/
PROVINCIAL**

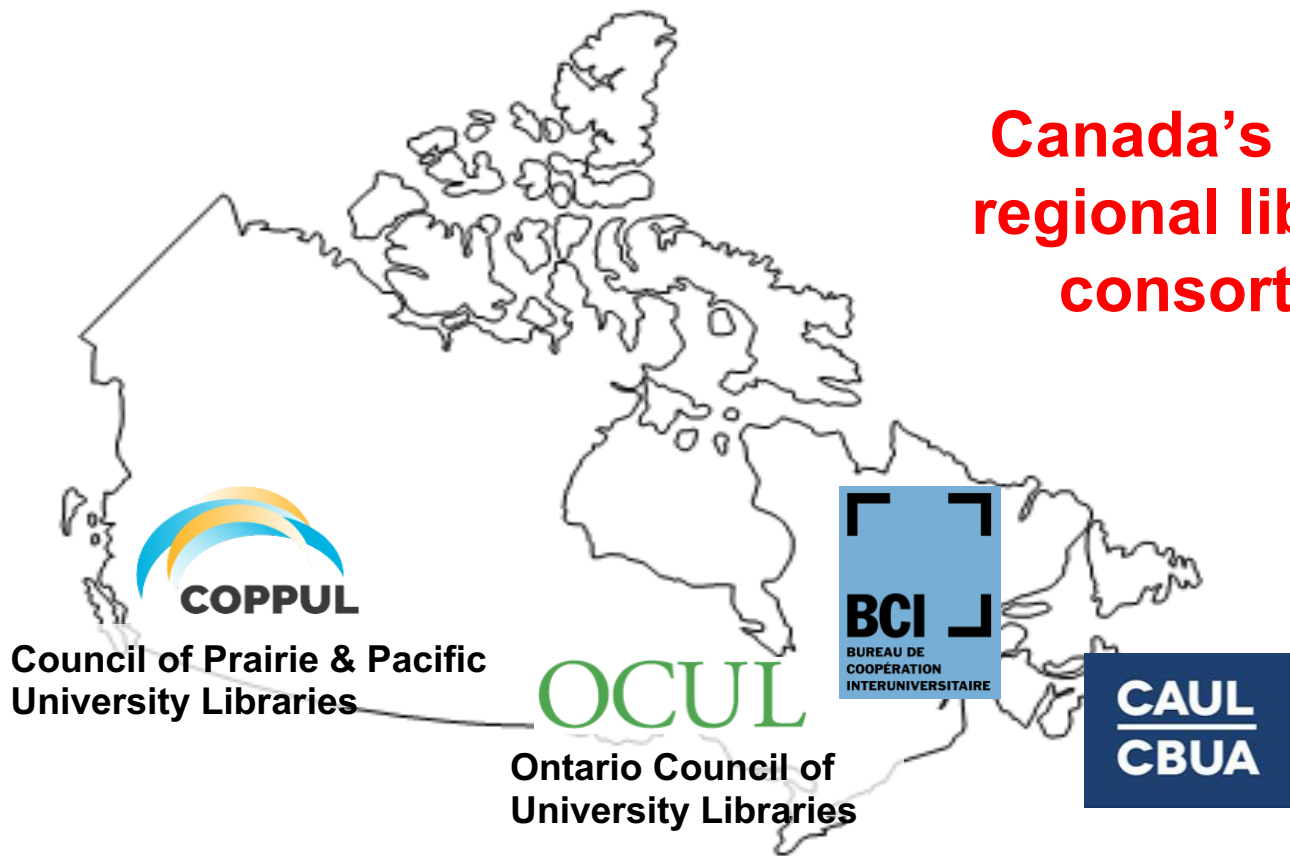
NATIONAL

Local Stakeholders

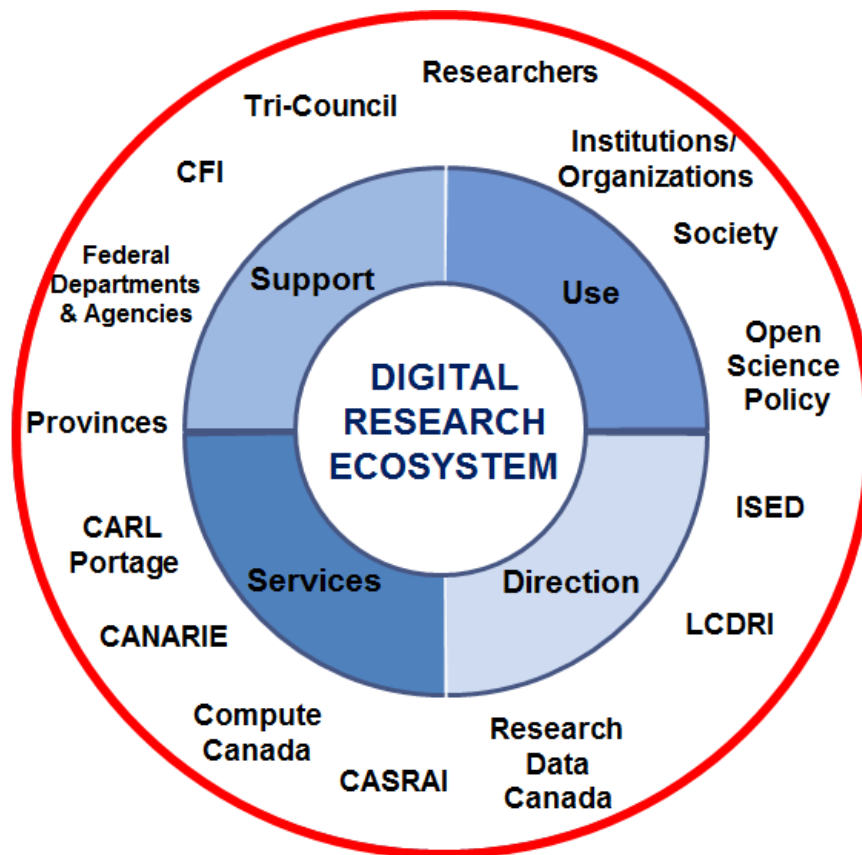


Regional Stakeholders

**Canada's four
regional library
consortia**



Digital Research Ecosystem



Goals of the Portage Network

- Build a **community of practice** for research data management (RDM)
- Engage and advocate for research data management with **stakeholder communities**
- Facilitate and provide leadership in the **development of RDM infrastructure**

RDM Infrastructure Platforms

Data Management Planning



Institutional
Strategy



DMPs



Deposit



Portage DMP Assistant

34 institutional accounts
2,662 registered users



Shared stewardship of research data

DMP Assistant is a bilingual tool for preparing data management plans (DMPs). The tool follows best practices in data stewardship and walks researchers step-by-step through key questions about data management.

- Step 1** Sign up with DMP Assistant.
- Step 2** Sign in and select a template under Organizations. The Portage template is the default.
- Step 3** Answer the questions that are relevant to your work. Guidance and examples are provided.
- Step 4** Revisit the tool throughout your research to review or revise your answers.

[Version française](#)

Template Covers

- Data collection
- Documentation & metadata
- Backup & storage
- Preservation
- Data sharing & reuse
- Responsibilities & resources
- Ethics & legal compliance

Portage DMP Assistant 2.0

Based on single, open-source platform: **DMP Roadmap**

Joint initiative:

- Digital Curation Centre
- University of California Curation Centre



New features:

- Provides foundation for machine actionable DMPs
- Supports DMPs as “living documents”
- Supports use of ORCID IDs
- Includes APIs for creating and sharing DMPs

SSHRC DMP Workshop

October 2016 - February 2017

Explored the use of DMPs by 15 SSHRC-funded research groups to identify related data management issues, barriers to uptake, and preparedness for widespread adoption.

Researchers completed the workshop with new perspectives on their research, and new practices to integrate into their workflows.



RDM Infrastructure Platforms

Data Management Planning



Institutional
Strategy



DMPs



Deposit



Research Data Repositories

Federated Research Data Repository



Dataverse North Working Group



Institutional
Strategy



DMPs



Deposit



Federated Research Data Repository

Feedback Log In Help EN



Find and Share Canadian Research Data

FRDR DFDR
FEDERATED RESEARCH DATA REPOSITORY DÉPÔT FÉDÉRÉ DE DONNÉES DE RECHERCHE

[Deposit Data](#) 

About FRDR

The Federated Research Data Repository (FRDR) is a joint initiative led by the Canadian Association of Research Libraries (CARL) and Compute Canada to provide Canadian researchers a place to deposit large data sets and to improve the discovery of Canadian research data.

[Learn more »](#)

Deposit Data

Any researcher affiliated with a Canadian institution can deposit data in the Federated Research Data Repository (FRDR) at no direct cost. The platform can efficiently move data sets of any size, and preservation and archival is done automatically. Research librarians from CARL curate and approve deposited items.

[Learn more »](#)

Find Data

Search the Federated Research Data Repository (FRDR) to find research data sets originating from researchers affiliated with Canadian institutions. Data deposited to other repositories across Canada can also be found by searching in FRDR. See the growing list of collaborating repositories by performing a search.

[Learn more »](#)

[Privacy Policy](#)

© 2017 Canadian Association of Research Libraries & Compute Canada



<https://portagenetwork.ca/frdr-dfdr>

Federated Research Data Repository

species at risk

Filter Results

8252 results found.

1 to 20 of 8252

Main Features:

- National **Discovery** Platform
- National **Repository** Option
- ‘Big Data’** upload/download
- Maintains **File Hierarchies**
- Preservation** Pipeline

Currently in “Limited Production”

<https://portagenetwork.ca/frdr-dfdr>

Limited Production Projects



Mountain Legacy Project:
Capturing Change in Canada's Mountains
U. Victoria

**Stone Tools, Diet, &
Sociality at the Dawn of Humanity**
U. Calgary

<https://www.facebook.com/OGSDS/photos/a.172692916509354.1073741828.148916232220356/354006195044691/?ty=pe=3&theater>
<https://www.facebook.com/BiblioArchives/photos/a.438306146206659.9019638503710620202/1301748839062381/?ty=pe=3&theater>
<https://www.usask.ca/water/>
<https://goo.gl/images/1Cw6fm>

Research Data Repositories



Main Features:

National **Discovery** Platform
National **Repository** Option
'Big Data' upload/download
Maintains **File Hierarchies**
Preservation Pipeline

Main Features:

Mature Platform
Widely adopted
User-friendly
Brandable

Dataverse North



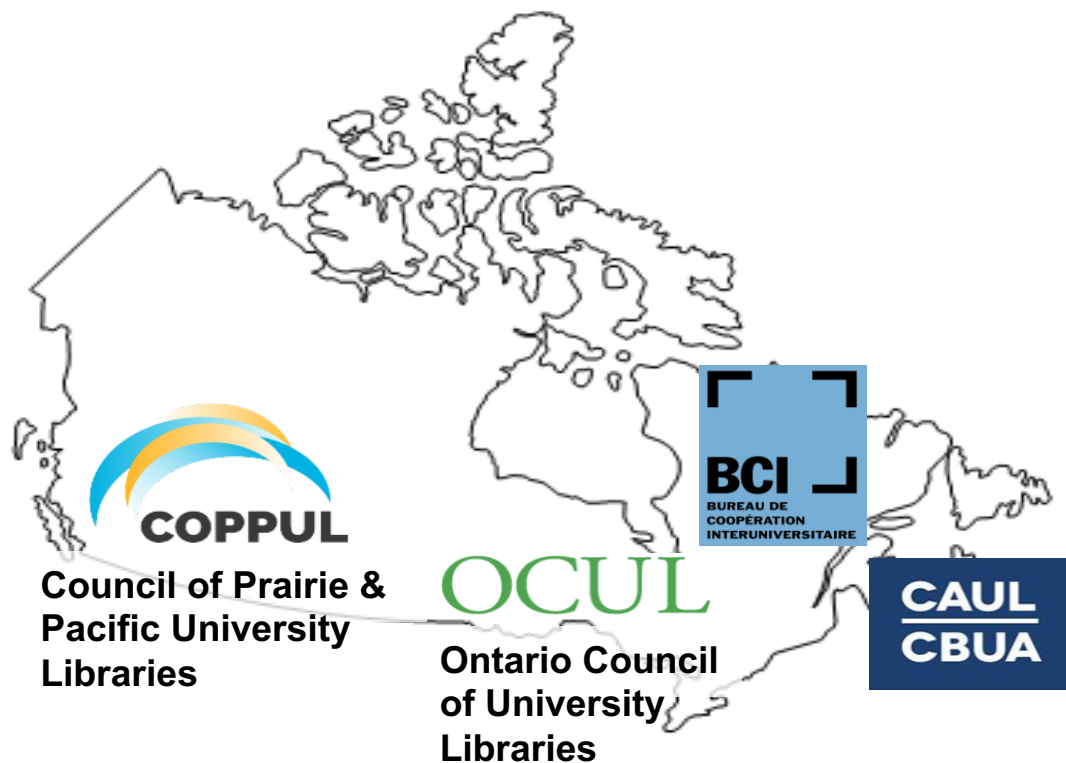
- A **community of practice** to coordinate training, support, outreach, and infrastructure development
- Explore **nationally coordinated** strategies and services

Dataverse North



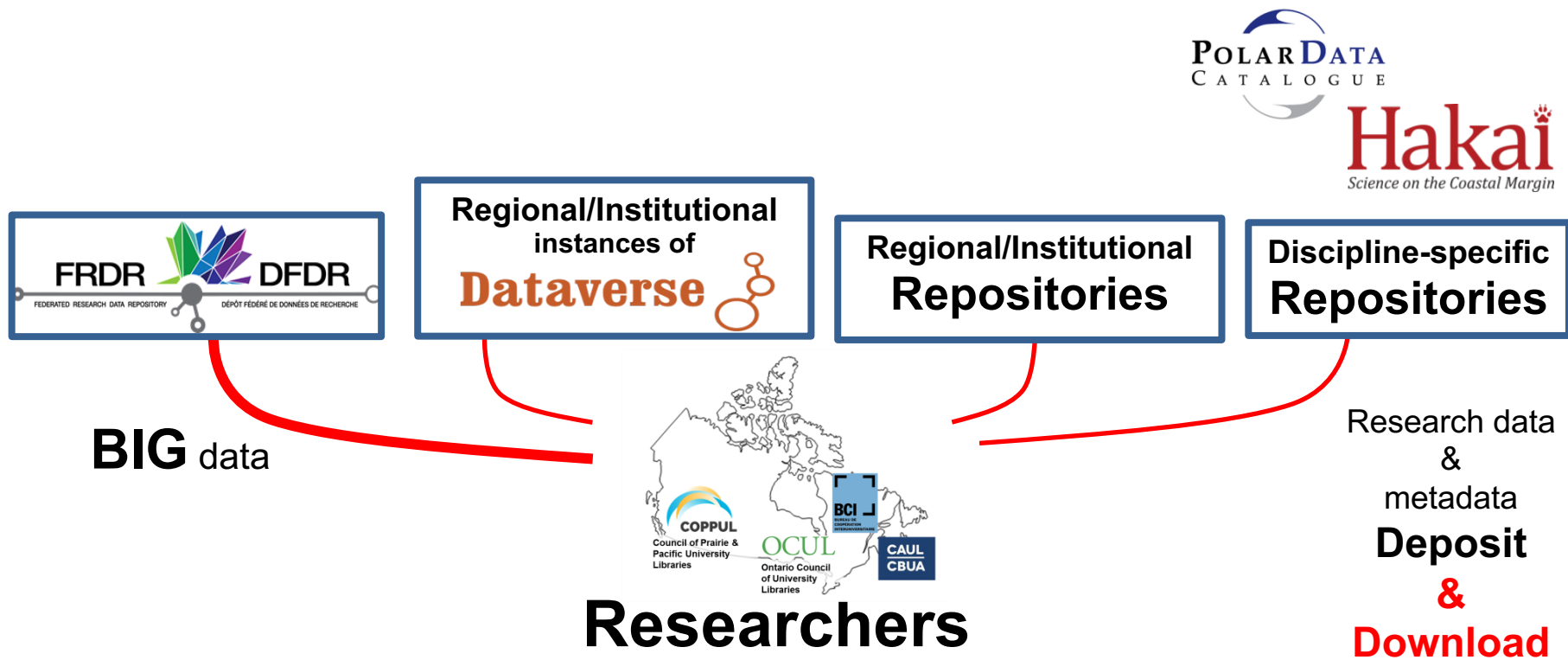
- Provide **equitable access** for Canadian researchers
- **Standardize services** across Canada
- **Support users** (librarians and researchers)
- Guide **feature developments and integrations**

Federated Architecture

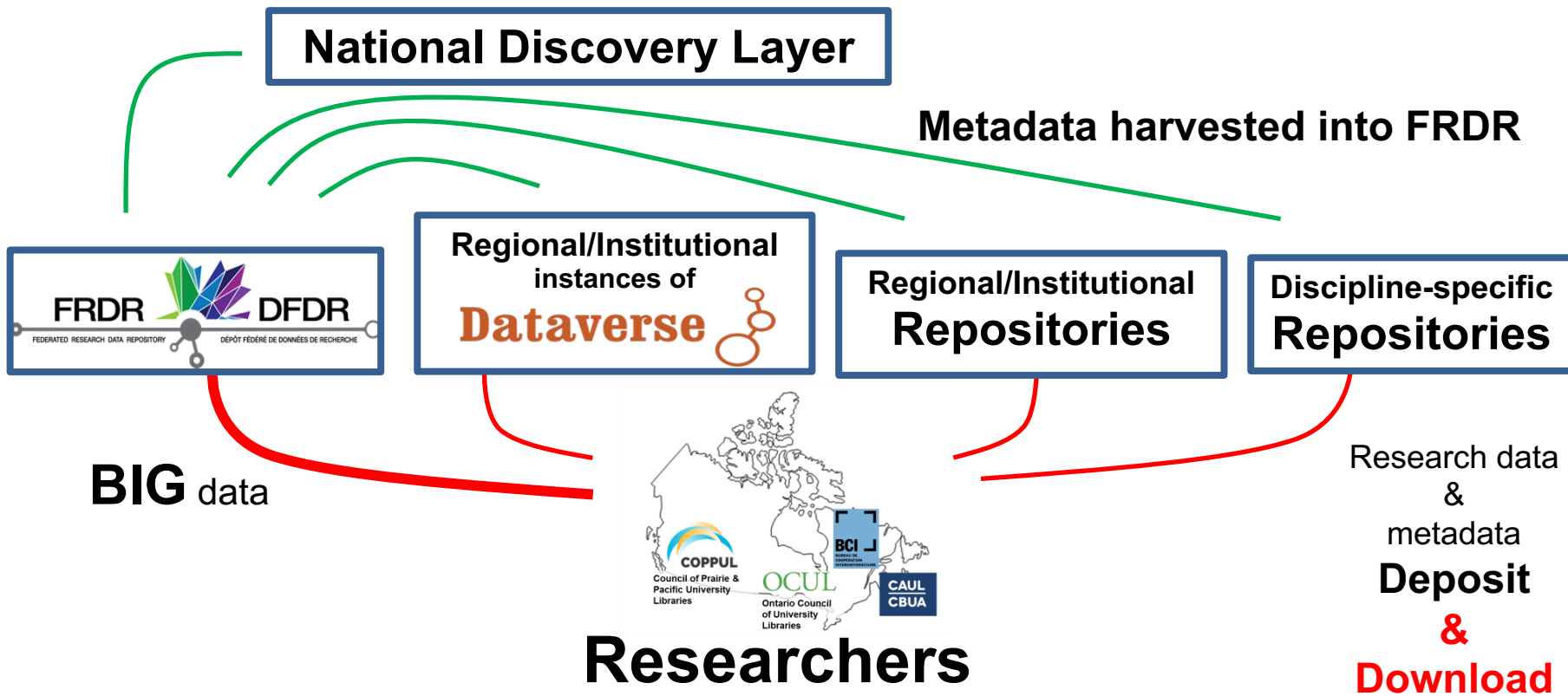


Researchers

Federated Architecture

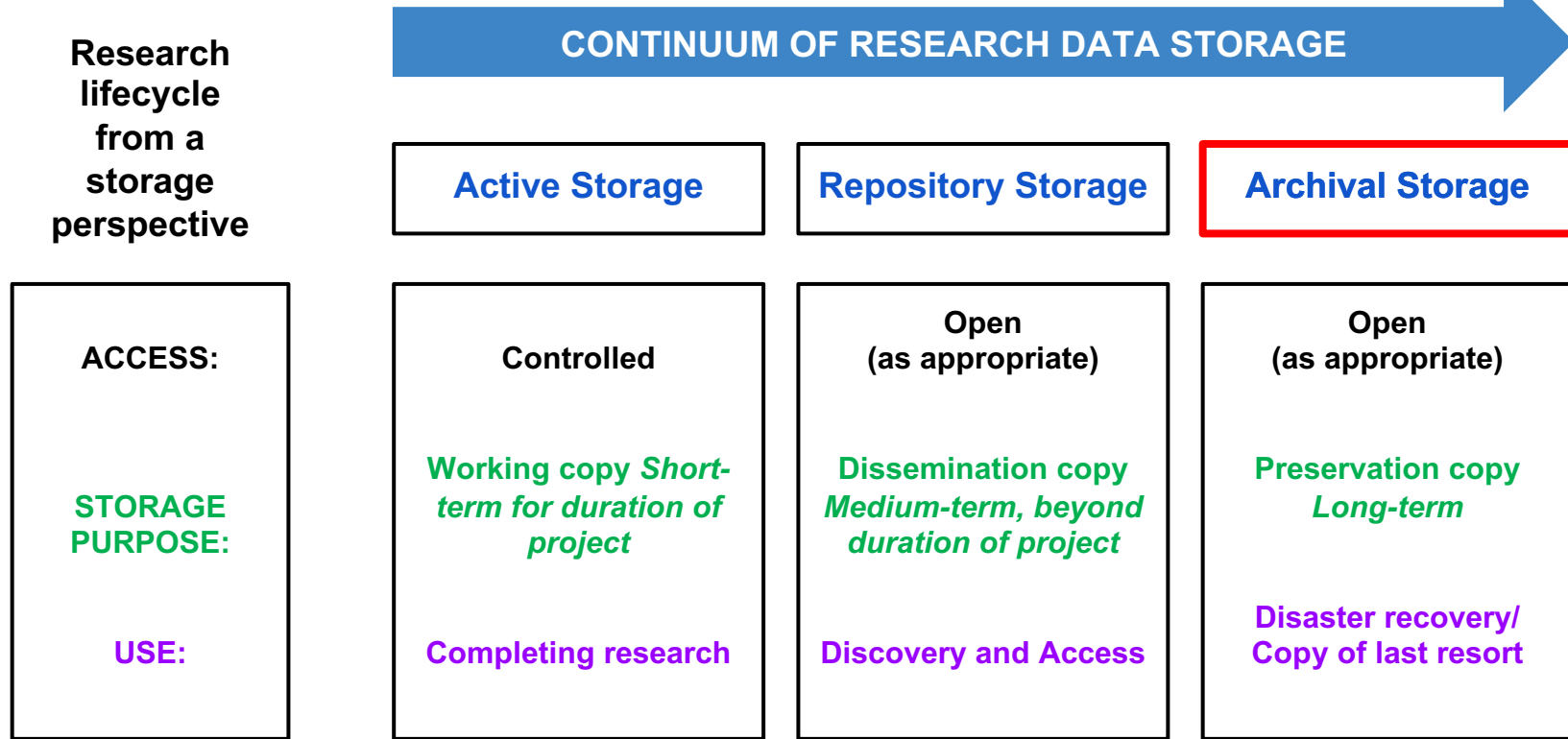


Federated Architecture



Data Storage Continuum

Preservation Expert Group White Paper
*****Coming Soon*****



Questions?



www.rdc-drc.ca
[@rdc_drc](https://twitter.com/rdc_drc)

Mark Leggott
Executive Director

mark.leggott@rdc-drc.ca



www.portagenetwork.ca
[@portageCARLABRC](https://twitter.com/portageCARLABRC)

Shahira Khair
CARL Project Officer

shahira.khair@carl-abrc.ca