portage

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

Canadian Leadership in RDM: Portage Partnerships & Progress

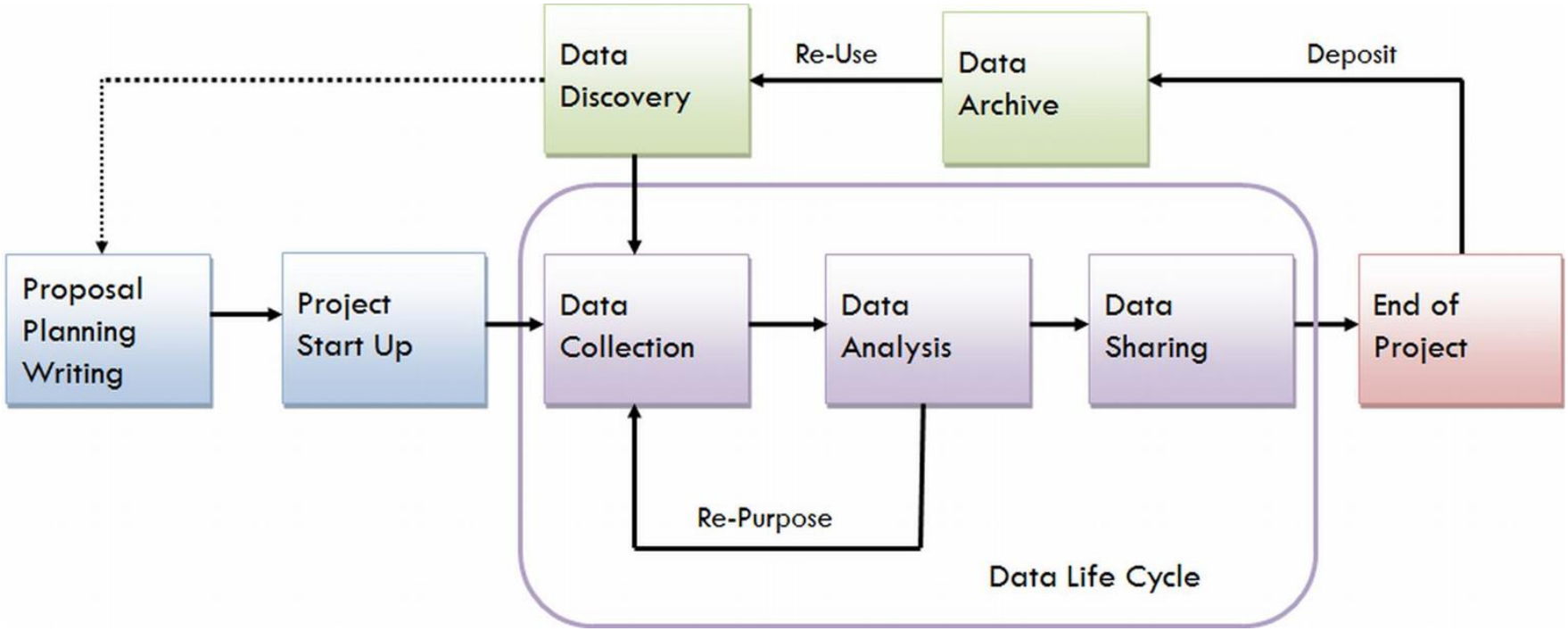
2018 McMaster Research Data Forum:
Common Challenges, Shared Solutions
Dec 12, 2018

Jeff Moon, Director, Portage

https://commons.wikimedia.org/wiki/File:Maple_Leaf.svg

What is Research Data Management?

A broad term used to describe the structure, organization, maintenance, and overall **stewardship** of research data.



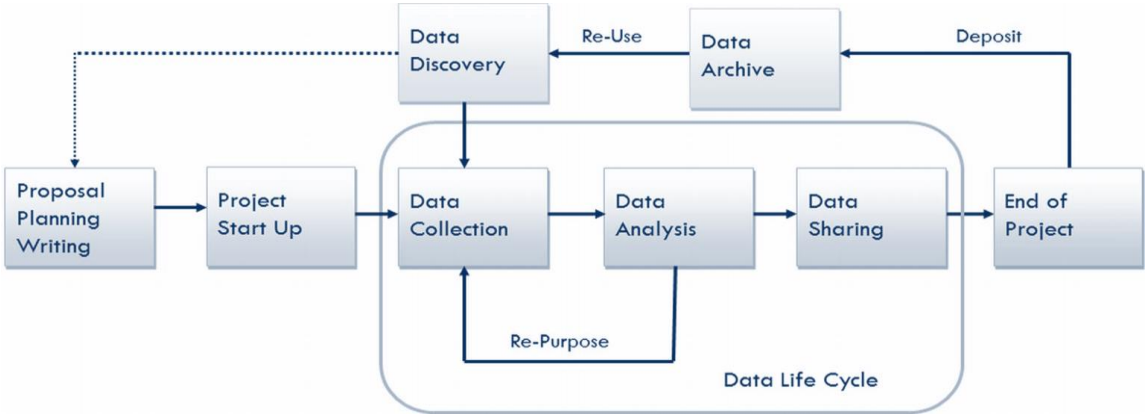
Research Lifecycle

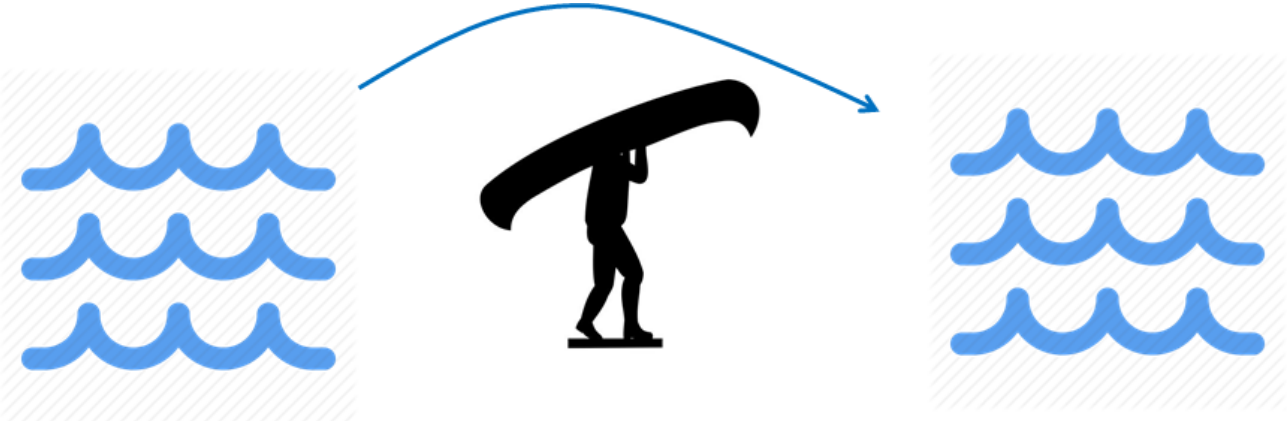
http://data.library.virginia.edu/data-management/lifecycle/

What is Portage?

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

"Portage is a CARL-sponsored national network of Research Data Management experts working collaboratively to build capacity and better coordinate RDM activities in Canada"





http://data.library.virginia.edu/data-management/lifecycle/

Outline

- Federal RDM policy developments
- Further context for RDM in Canada
- Goals and outputs of Portage
- Developing a shared vision for RDM in Canada

Outline

- **Federal RDM policy developments**
- Further context for RDM in Canada
- Goals and outputs of Portage
- Developing a shared vision for RDM in Canada

Research Data Management – Recent Policy Developments

CIHR IRSC
Canadian Institutes of Health Research
Instituts de recherche en santé du Canada

NSERC
CRSNG

SSHRC CRSH

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

Research Excellence
Research Dissemination
Research Impact
Research Best Practice

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

Research Excellence
Research Dissemination
Research Impact
Research Best Practice

Project Management

Reproducibility

Avoiding Duplication

Meeting
Journal Guidelines

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

Research Excellence

Research Dissemination

Research Impact

Research Best Practice

Data Sharing

Citation

Interdisciplinarity

OPEN ACCESS Freely available online

PLoS one

Sharing Detailed Research Data Is Associated with Increased Citation Rate

Heather A. Piwowar*, Roger S. Day, Douglas B. Fridsma

Department of Biomedical Informatics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, United States of America

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

Research Excellence
Research Dissemination
Research Impact
Research Best Practice

Within Science

Within Society

In Policy Development

http://www.ands.org.au/news-and-events/dataimpact

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

Research Excellence
Research Dissemination
Research Impact

Research Best Practice

Transparency

Trust

Ethics

Responsible Use of Public Funds

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

- **2013**
- Capitalizing on Big Data: Toward a Policy Framework for Advancing Digital Scholarship in Canada

- **2016**
- Tri-Agency Statement of Principles on Digital Data Management

- **2017-2018**
- Draft Tri-Agency Research Data Management Policy

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

Draft Tri-Agency Research Data Management Policy

Online Consultations

OPENED: June 5, 2018

CLOSED: Sept 30, 2018

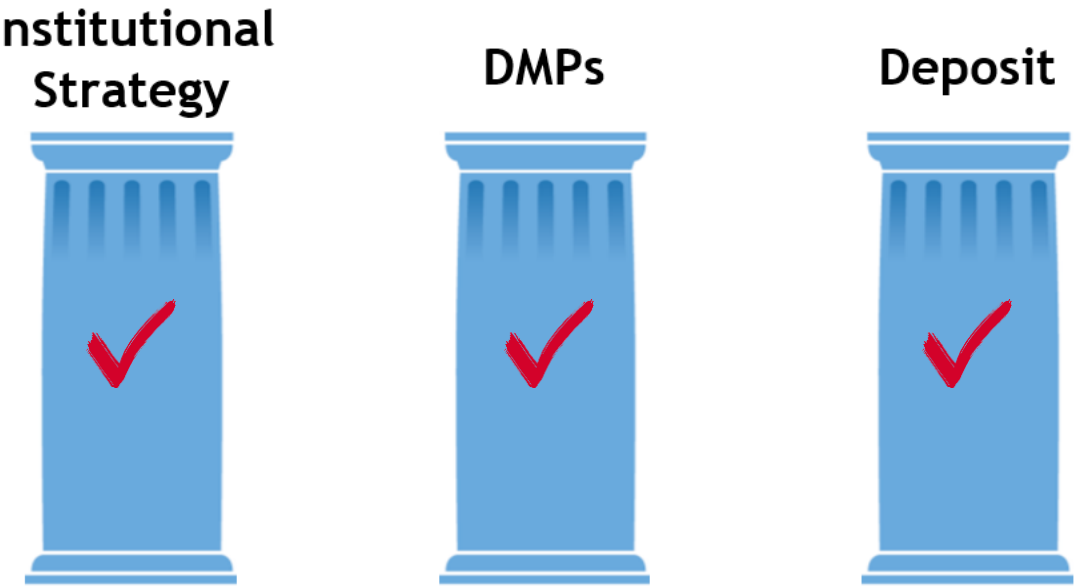
SSHRC: ResearchData-Donneesderecherche@sshrc-crsh.gc.ca

NSERC: ResearchData-Donneesderecherche@nserc-crsng.gc.ca

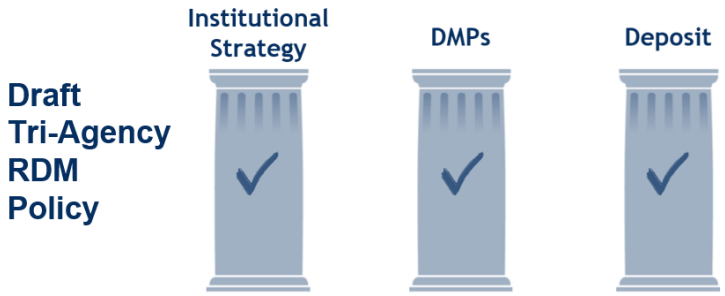
CIHR: ResearchData-Donneesderecherche@cihr-irsc.gc.ca

Research Data Management – Recent Policy Developments

Draft Tri-Agency RDM Policy



Research Data Management – Recent Policy Developments



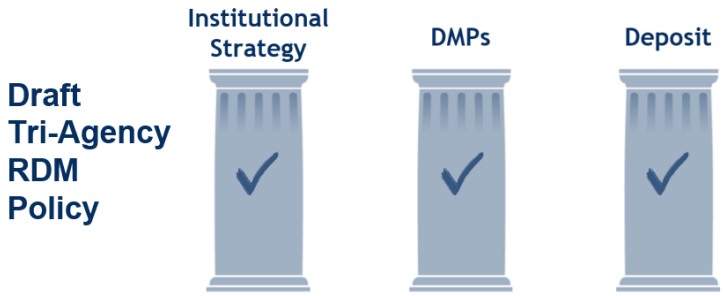
1. Institutions: Institutional Strategy

- Institution administering Tri-Agency funds would need an **institutional research data management strategy** outlining how the institution will provide its researchers with an environment that enables and supports world-class RDM practices.

- Strategies would be posted on the institution's website, with contact information to direct inquiries about the strategy.

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

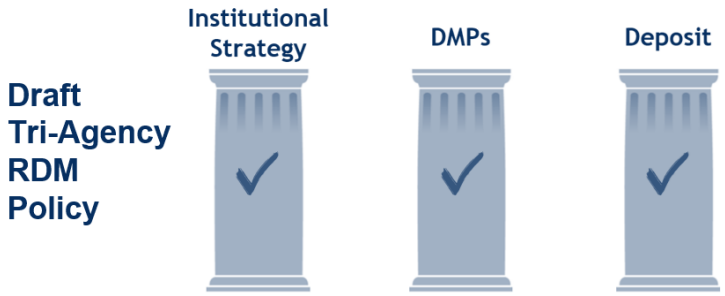


Why Require Institutional Strategies?

- Recognizes the role of institutions in RDM
- Opportunity to identify and address gaps at campus level
- Opportunity to investigate local solutions while leveraging regional, provincial, and national services and platforms
- Helps inform Tri-Agencies on institutional data management capacity
- Serves as foundation for potential requirements that follow

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

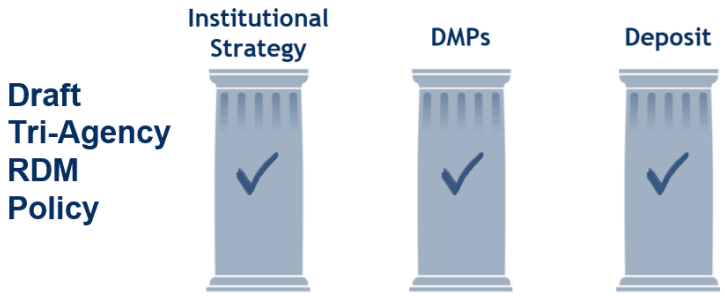


2. Researchers: Data Management Plans

- **DMPs required for research projects** supported wholly or in part by Tri-Agency funds. Submission of DMPs would be a condition of the release of grant funds.
- *For specific funding opportunities*, DMPs would be **submitted at time of application** to the appropriate agency, for possible consideration in the adjudication process.

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

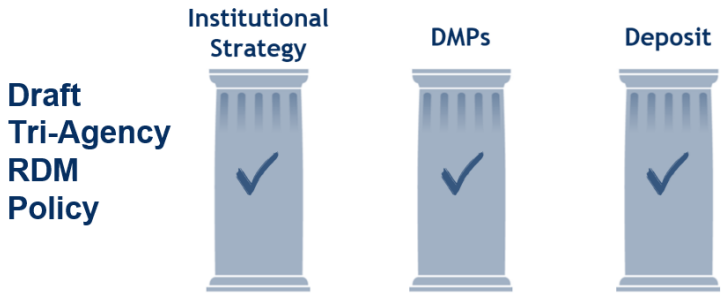


Why Require Data Management Plans?

- DMPs are an emerging international best practice
- DMPs help researchers anticipate RDM opportunities and challenges early in the research life cycle
- Researchers report that developing DMPs helps improve research processes, methodologies, and outcomes
- DMPs could feed into research ethics approval processes
- DMPs help anticipate issues related to ownership of data, potential for data sharing, etc.

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

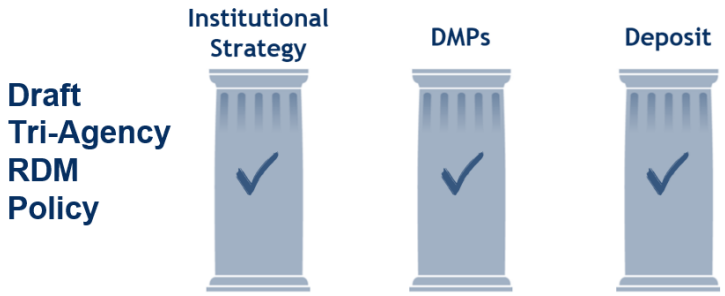


3. Researchers: Data Deposit

- Research **data**, **metadata**, and **code** supporting journal publications, pre-prints and other research outputs arising from funded research would be **deposited in an appropriate public repository** or other platform that will ensure safe storage, preservation, curation, and (if applicable) access to the data.

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments



Why Require Data Deposit?

- Data, metadata, and code safe & **available to data creator** & **appropriately to others** in future
- Ensure **proper use of public funds**
- Facilitate **reproducibility**
- Facilitate appropriate **data sharing**

Slides derived from Tri-Agency presentation, April 19, 2018

Research Data Management – Recent Policy Developments

More on Policy Requirements later...

Outline

- Federal RDM policy developments
- **Further context for RDM in Canada**
- Goals and outputs of Portage
- Developing a shared vision for RDM in Canada

Other factors making Research Data Management important…

Growing public awareness of data

Recognition of the value of data in advancing research

Addressing pressures from commercial concerns

Need to verify research results

RDM

Government policy directions

Growing demand for data

Journal policies

https://www.iconsdb.com/red-icons/question-mark-icon.html

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

Growing public awareness… RDM in the news

CBC

> *"Now what we're getting is massive amounts of data and really the biggest challenges we have are data management."*
>
> - Mike Suitor

http://www.cbc.ca/radio/spark/385-online-dating-genetic-data-and-more-1.4538832/caribou-herds-get-new-wearable-tech-1.4538840

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

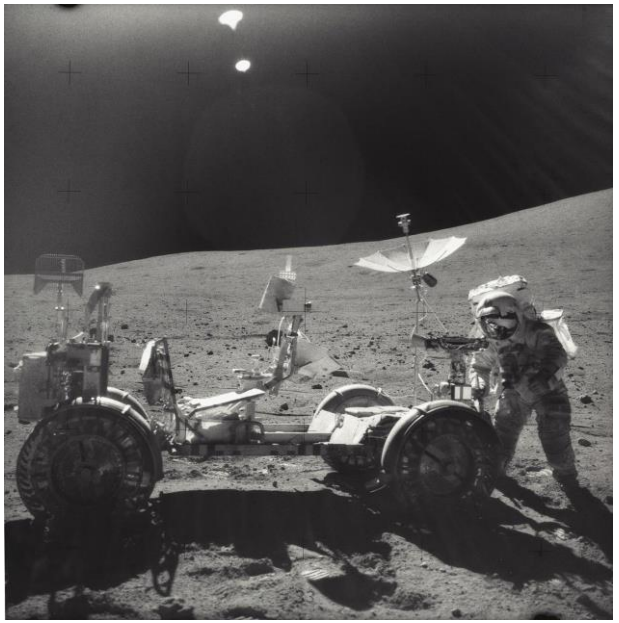
Growing public awareness… RDM in the news

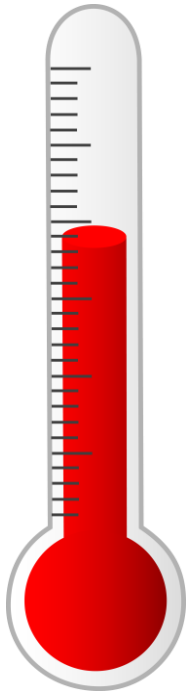
Live Science > Space

'Lost' NASA Tapes Show Humans Sort of Caused Global Warming on the Moon Too

By Brandon Specktor, Senior Writer | June 13, 2018 06:05am ET

'…hundreds of the reels of magnetic tape holding those records went missing nearly 40 years ago, thanks to an archival blunder'





JOURNAL OF GEOPHYSICAL RESEARCH
Planets
AN AGU JOURNAL

Research Article | Full Access

Examination of the Long-Term Subsurface Warming Observed at the Apollo 15 and 17 Sites Utilizing the Newly Restored Heat Flow Experiment Data From 1975 to 1977

S. Nagihara, W. S. Kiefer, P. T. Taylor, D. R. Williams, Y. Nakamura

First published: 25 April 2018 | https://doi.org/10.1029/2018JE005579

https://geschichte.inf.tu-dresden.de/?q=EN/NORMAL/expo06/content
http://www.clipartpanda.com/clipart_images/red-thermometer-clip-art-8329215
https://www.livescience.com/62805-moon-surface-warming-apollo-astronauts.html

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

Recognition of the value of data: *within a discipline*

Practice and Research in Nursing (PRN) Wound Studies, 1999-2009 [Canada]

"Many thanks for sharing the raw data. If more research groups and industry would do this we could save more lives and limbs!"

Ronald J. Shannon, MPH
President, Global Health Economic Projects, LLC
56 Via Da Vinci, Clifton Park, NY 12065

Scholars Portal Dataverse

Queen's University Library Research Data Archive Dataverse
https://hdl.handle.net/10864/CORX8

Consensus opinion on compression therapy and venous ulcers
Welsh, Lynn. **Primary Health Care (2014+); London** Vol. 26, Iss. 4, (Apr 2016): 26.
DOI:10.7748/phc.26.4.26.s28

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

Recognition of the value of data: *between disciplines*

Five Weird Archives That Scientists Use to Study Past Climates

When tree rings, ice cores, and cave formations can't cut it, try your luck with whale earwax or bat poop.

Data collected on
tree rings
ice cores
cave formations
whale earwax &
bat poop

& sediment in ancient Roman aqueducts

now being used by **climate scientists** to study global warming

Citation: Wendel, J., and M. Kumar (2018), Five weird archives that scientists use to study past climates, *Eos, 99,* https://doi.org/10.1029/2018EO095483. Published on 30 March 2018.

Thanks to **Kevin Lindstrom,** Physical Sciences and Engineering Librarian, Woodward Library,2198 Health Sciences Mall
University of British Columbia, for pointing out this citation

The need to *verify* research results

Why scientists need to do more about research fraud

Richard P Grant

Scientific misconduct is more than just an academic problem – it has repercussions for real people

https://www.theguardian.com/science/occams-corner/2018/jan/04/science-fraud-research-misconduct

Addressing pressures from commercial concerns

SPRINGER NATURE
Research Data Support

- *Advice on finding repositories*
- *Advice on complying with policies*
- *Check for presence of sensitive information*
- *Advice on data availability statements/data citation*
- *Ensure accessibility of files for end users*
- *Easy upload to Springer Nature Figshare*
- *Ensure files match metadata & assoc. pubs*
- *Organize files into logical structure & collections*
- *Match authors to associated publications*
- *Copy-edit submitted metadata*
- *Generate Digital Object Identifier (DOI)*
- *Categorize content according to industry standard definitions*
- *Enable embargoes with private, anonymous access enabled*
- *Link and synchronize dataset release to associated publications*
- *Enhancement report and feedback from Research Data Editor*

http://www.exlibrisgroup.com/wp-content/uploads/2018/02/ALA-MW-2018-Soderdahl-Grant-Neuwirth-Esploro-2-9.pdf
https://www.elsevier.com/about/press-releases/science-and-technology/elsevier-launches-mendeley-data-to-manage-entire-lifecycle-of-research-data

Addressing pressures from commercial concerns

SPRINGER NATURE
Research Data Support

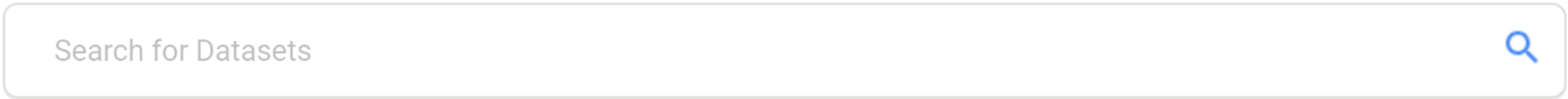
- *Advice on finding repositories*
- *Advice on complying with policies*
- *Check for presence of sensitive information*
- *Advice on data availability statements/data citation*

"My complaint is **they're "solving" the 'easy' RDM problems**, **for a fee**"

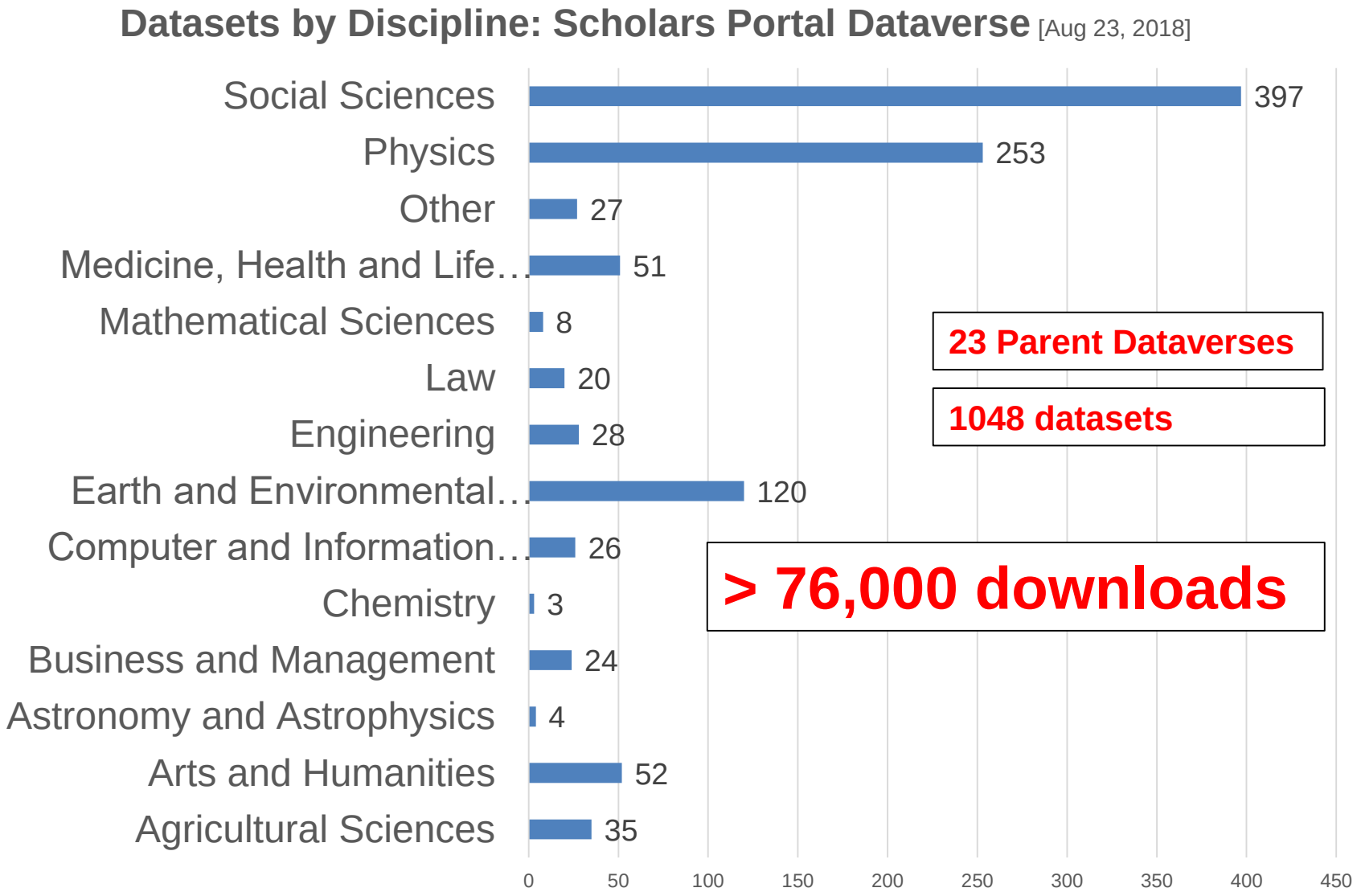
- *Copy-edit submitted metadata*
- *Generate Digital Object Identifier (DOI)*
- *Categorize content according to industry standard definitions*
- *Enable embargoes with private, anonymous access enabled*
- *Link and synchronize dataset release to associated publications*
- *Enhancement report and feedback from Research Data Editor*

http://www.exlibrisgroup.com/wp-content/uploads/2018/02/ALA-MW-2018-Soderdahl-Grant-Neuwirth-Esploro-2-9.pdf
https://www.elsevier.com/about/press-releases/science-and-technology/elsevier-launches-mendeley-data-to-manage-entire-lifecycle-of-research-data

And we can now include Google in the mix…

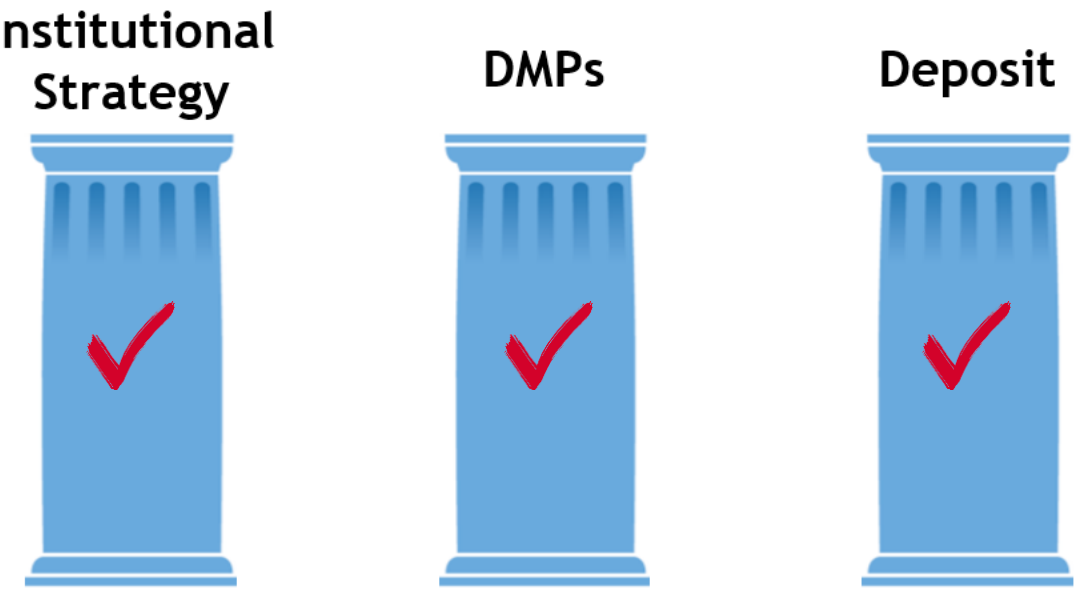


Growing demand for data



Government policy directions

Emerging Tri-Agency RDM Policy



Drivers also include:

- Trends in **journal publishing**
- Movement toward **open science**
- International initiatives & best practices

Government funding – 2018 Canadian Budget

Harnessing Big Data

"The Government proposes to provide **$572.5 million** over five years, with $52 million per year ongoing, to implement a **Digital Research Infrastructure Strategy** that will deliver more open and equitable access to advanced computing and big data resources to researchers across Canada."

p. 92 https://www.budget.gc.ca/2018/docs/plan/budget-2018-en.pdf

Definition of ‘Digital Research Infrastructure’

Industry Canada Industrie Canada

“DRI refers to the elements required to perform data-intensive and computationally-intensive research and data management, including high-performance computing, storage, high-speed networks and other tools and resources, including software, standards and data management services.” p. 3

https://www.ic.gc.ca/eic/site/icgc.nsf/vwapj/DRI_Strategy_Consultation-eng.pdf/$file/DRI_Strategy_Consultation-eng.pdf

Definition of ‘Digital Research Infrastructure’

Industry Canada Industrie Canada

“DRI refers to the elements required to perform data-intensive and computationally-intensive research and **data management**, including high-performance computing, storage, high-speed networks and other tools and resources, including software, standards and **data management services**.” p. 3

https://www.ic.gc.ca/eic/site/icgc.nsf/vwapj/DRI_Strategy_Consultation-eng.pdf/$file/DRI_Strategy_Consultation-eng.pdf

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

Outline

- Federal RDM policy developments
- Further context for RDM in Canada
- **Goals and outputs of Portage**
- Developing a shared vision for RDM in Canada

Portage has many partners in RDM…

compute canada | calcul canada

CARL ABRC

portage

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

RDC DRC

Research Data Canada – Données de recherche Canada

OCUL Ontario Council of University Libraries

Scholars Portal

A Service of the Ontario Council of University Libraries

UNIVERSITY OF ALBERTA

CIHR IRSC

Canadian Institutes of Health Research | Instituts de recherche en santé du Canada

NSERC CRSNG

SSHRC CRSH

… and many other RDM and DRI stakeholders

https://static1.squarespace.com/static/576bfb4b2e69cfe8cbce5a20/5786be9eff7c509730365440/578a66eabebafb8c81408a37/1468688108359/172+-+Canoe+Portage+Sign+T-Shirt+-+Canada+Clothing.jpg?format=500w
https://cdn0.iconfinder.com/data/icons/wthr-color/32/693734-tide-512.png

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**

Portage origins and progress…

New Business Models

CIHR Training Modules

SSHRC DMP Workshop

COPPUL MOU

CSA -Life Sciences

Website Launch

Compute Canada/FRDR

DMP Assistant

First Portage Director

CARL Approval

2016

15

2015

2014: Project ARC

DVN Report

BCI MOU

Sensitive Data WG

Regional Library Consortia

Portage Day - Toronto

Institutional Strategy WG

Second Portage Director

FRDR – Limited Production

Portage Day - Montreal

Formative Assessment

LCDRI

CAUL MOU

Dataverse North

OCUL MOU

New Governance Model

115

2018

98

2017

42

Colleges

Non-CARL Libraries

International Outreach

Portage/Tri-Agency Tour

Preservation White Paper

CARL Directors Support

2019

2018

Governance

Platforms/Services

Outreach

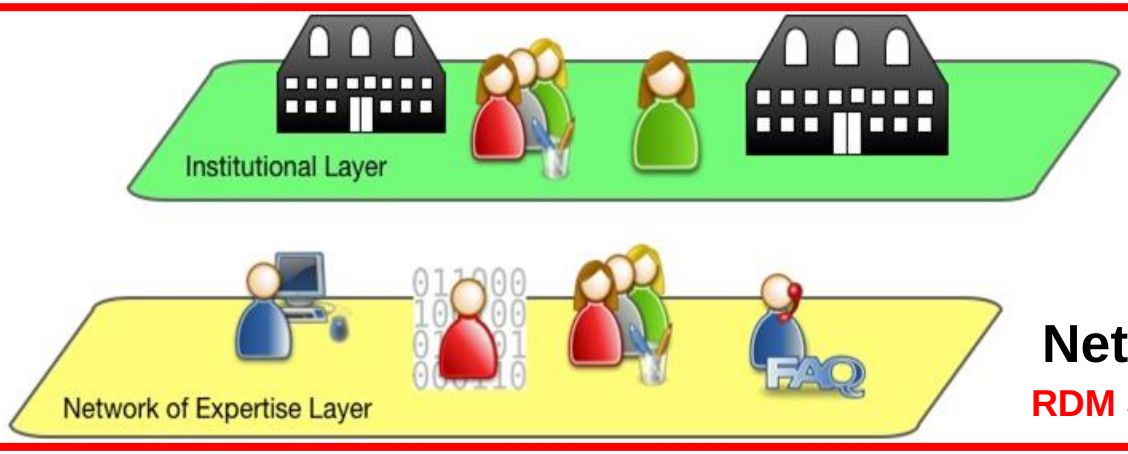
Agreements

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

Goals of the Portage Network

- Build a **community of practice** for research data management (RDM)

Building a Community of Practice



Institutions

Network of Experts

RDM Specialists/Librarians

Expert Groups

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

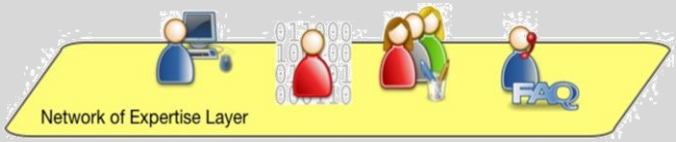
Working Groups

- Dataverse North WG
- Responsible Research Data Management Practices for Sensitive Data
- FRDR Service Model WG
- FRDR Discovery Service WG
- Institutional RDM Strategies WG

➢ ***110+ Experts from across Canada***

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

Broad representation across Canada



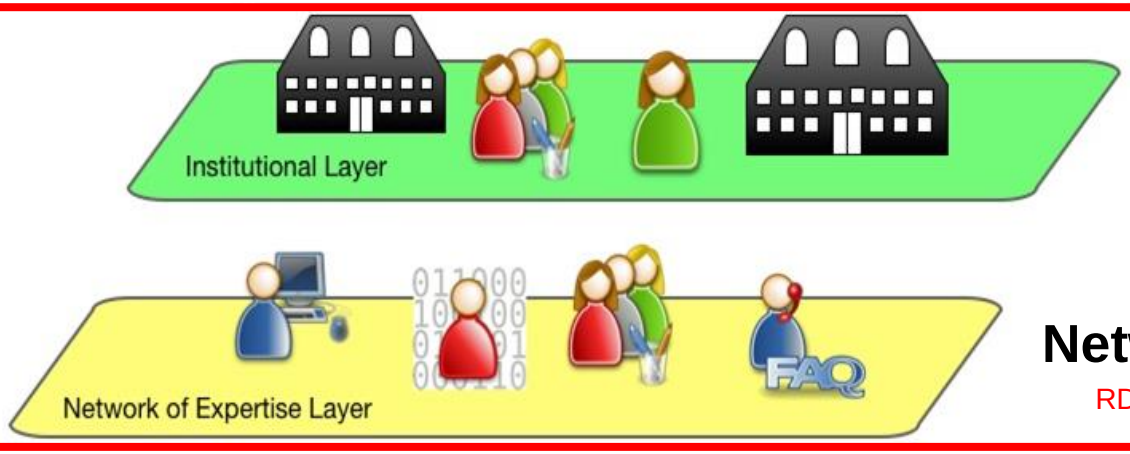
22 CARL Library Institutions

Brock University	Acadia University
Carleton University	Cape Breton University
Dalhousie University	Concordia University
McGill University	HEC Montréal
McMaster University	Mount Saint Vincent University
Queen's University	Saint Mary's University
Ryerson University	Trent University
Simon Fraser University	University of Notre Dame
Université de Montréal	University of Prince Edward Island
University of Alberta	Wilfrid Laurier University
University of British Columbia	AceNet
University of Calgary	Canadian Institutes of Health Research
University of Manitoba	CAREB
University of New Brunswick	CASRAI-CA
University of Ottawa	Compute Canada
University of Saskatchewan	Council of Prairie and Pacific University Libraries
University of Toronto	Cybera
University of Waterloo	First Nations Information Governance Centre
University of Western Ontario	Métis Nation Council
University of Windsor	National Research Council Canada
York University	Research Data Canada
Library and Archives Canada	Scholar's Portal
	Secretariat on Responsible Conduct of Research
	Statistics Canada
	BCI
	CARL

26 Other Institutions & Organizations

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

Building a Community of Practice



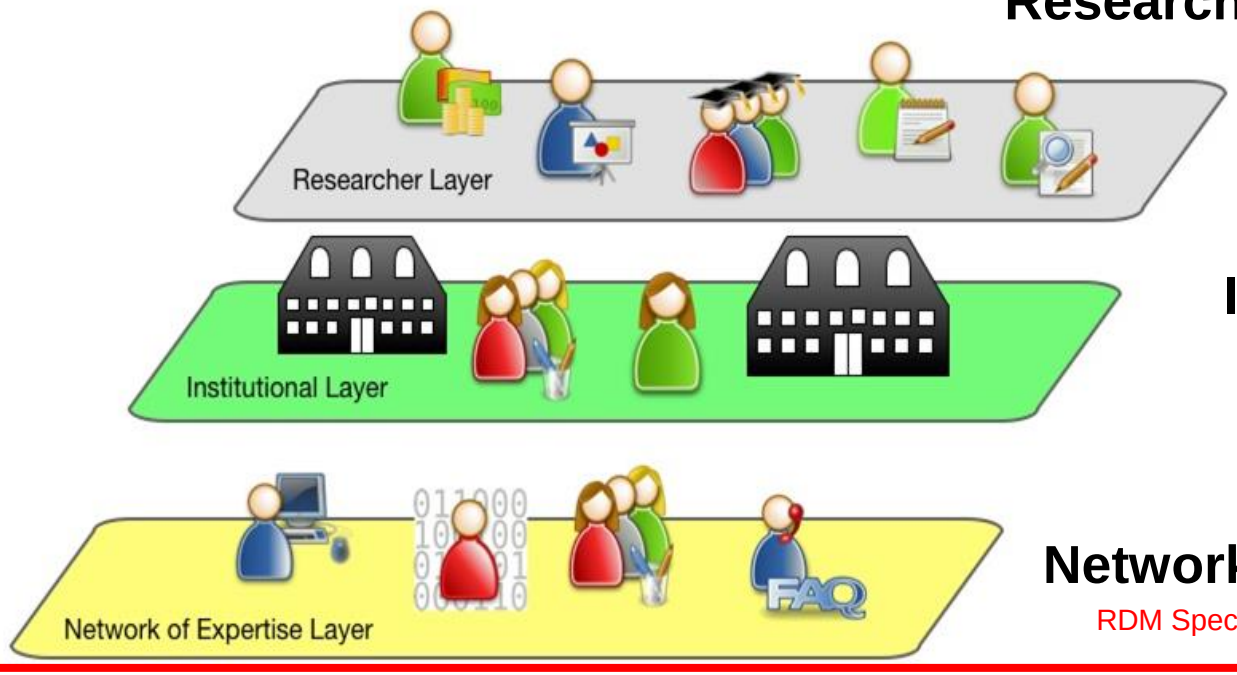
Institutions

Liaison Librarians
IT Specialists
Research Offices
Ethics Boards
Other stakeholders

Network of Experts

RDM Specialists/Librarians

Building a Community of Practice



Researchers

Academic
Government
Others

Institutions

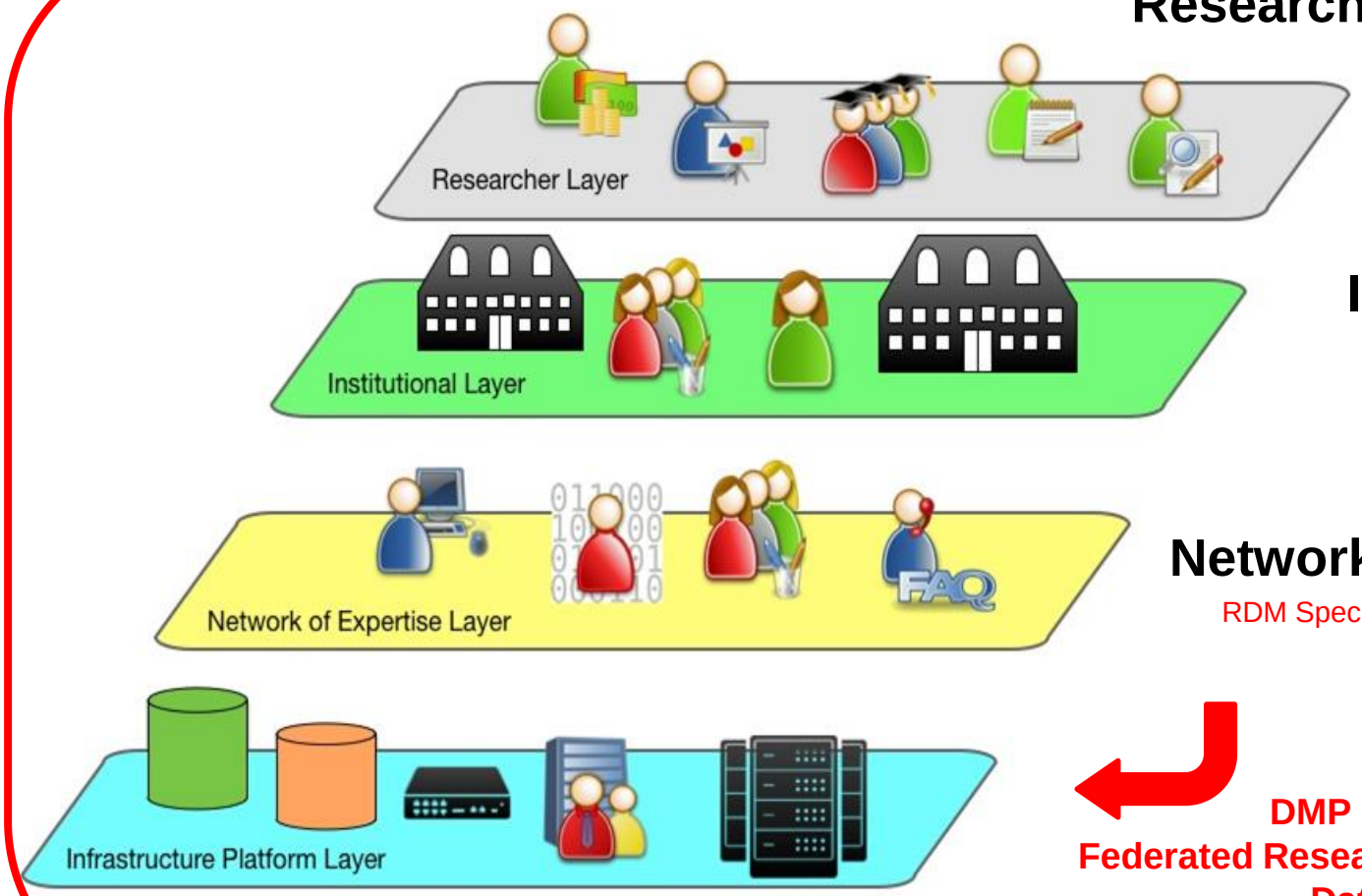
Liaison Librarians
IT Specialists
Research Offices
Ethics Boards
Other stakeholders

Network of Experts

RDM Specialists/Librarians

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

Building a Community of Practice



Researchers

Academic
Government
Others

Institutions

Liaison Librarians
IT Specialists
Research Offices
Ethics Boards
Other stakeholders

Network of Experts

RDM Specialists/Librarians

DMP Assistant
Federated Research Data Repository
Dataverse

Infrastructure Platforms

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

Network of Experts Publications

Selected Port

- **Portage Netwo** Graduate Stud
- **Portage Progr** Research Libra
- **The SSHRC DM** Director of the
- **Towards a Coll** paper presente Humphrey, Kat (PDF)
- **Multi-Stakehol** presented at IA Bourne-Tyson,
- **Research Data** (Presentation),

Strategic Doc

- **Formative Ass** 2017 (PDF)
- **Portage Business Plan:** 2017 and 2018 (PDF)
- **Portage: Organizational Framework**, April 2015 (PDF)

CARL ABRC
CANADIAN ASSOCIATION OF RESEARCH LIBRARIES · ASSOCIATION DES BIBLIOTHÈQUES DE RECHERCHE DU CANADA

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

Portage Comments on
Draft Tri-Agency Research Data Management Policy

Original Tri-Agency DRAFT POLICY & FAQ

Background

The Portage Network, an initiative of the Canadian Association of Research Libraries (CARL), is dedicated to the advancement of research data management (RDM) in Canada. Portage, through its nation-wide network of over 115 experts from over 40 organizations, works proactively with community partners to develop national RDM services and infrastructure in support of all researchers in Canada.

Portage had the privilege of working closely with Tri-Agency representatives during the pre-consultation phase of policy development, traveling the country to meet with 14 different stakeholder groups in 4 cities in Jan-Feb 2018. These meetings provided an ideal opportunity both to share Tri-Agency policy requirements with researchers and to describe Portage services and infrastructure designed to help researchers fulfill each requirement. Additionally, these meetings gave Portage, alongside our Tri-Agency colleagues, the chance to hear feedback from stakeholder groups and to respond together to questions that arose.

Portage also took the opportunity to convene three Town Hall meetings in July 2018, engaging over 32 participants. Feedback from these Town Halls and other less formal discussions was considered when crafting the Portage response provided below. A summary of comments from the community is provided in **Appendix A**.

ave
tive
ey
ing
a

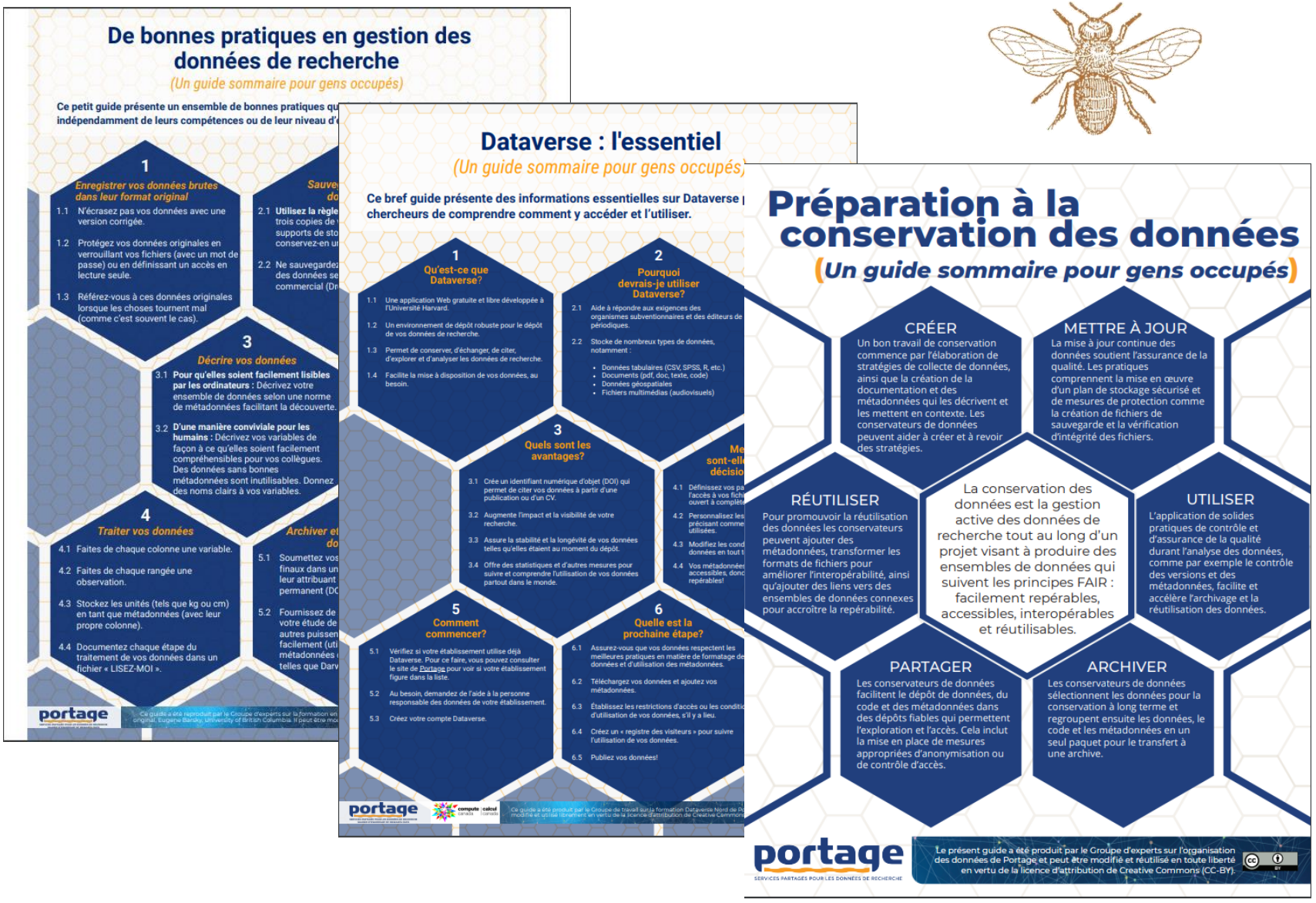
-

Management

- **Developing a Digital Research Infrastructure Strategy**, September 2015 (PDF)
- **CARL Response to the Draft Tri-Agency Statement of Principles on Digital Data Management**, September 2015 (PDF)

Busy Bees – *brief guides for busy people*

Busy Bees – *aussi disponible en français*



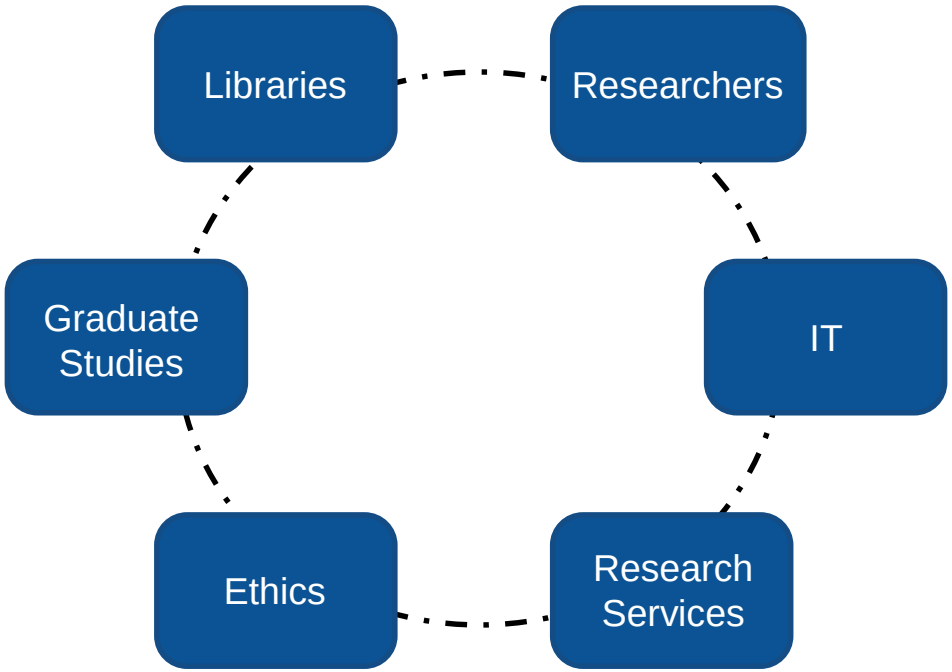
Goals of the Portage Network

- Build a **community of practice** for research data management (RDM)

- Engage and advocate for research data management with **stakeholder communities**

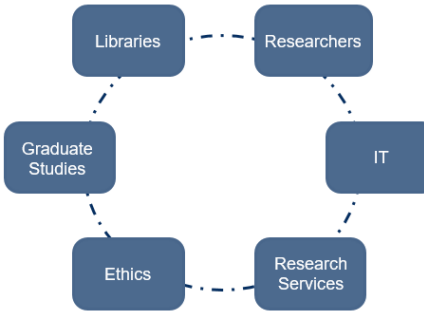
Levels of Stakeholder Engagement

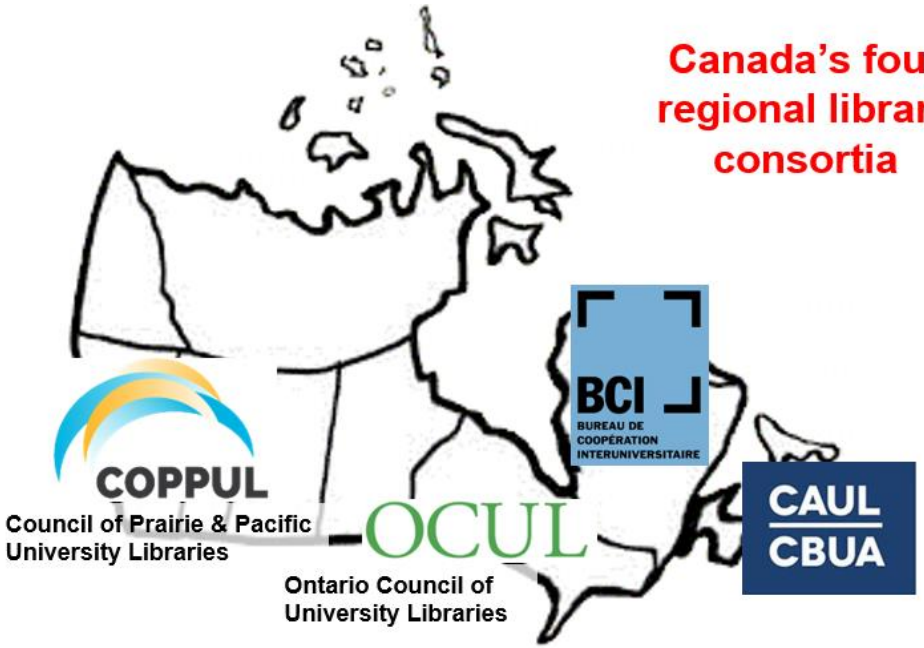
LOCAL | **REGIONAL/ PROVINCIAL** | **NATIONAL**



Levels of Stakeholder Engagement

LOCAL | **REGIONAL/ PROVINCIAL** | **NATIONAL**





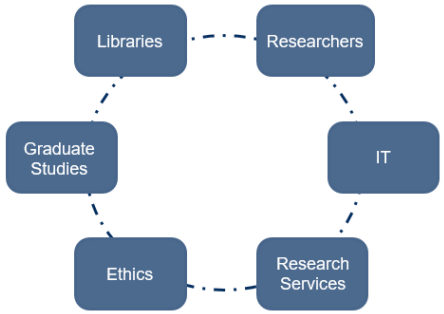
portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

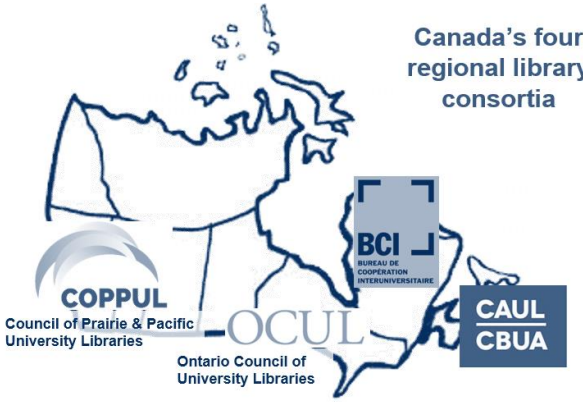
Levels of Stakeholder Engagement

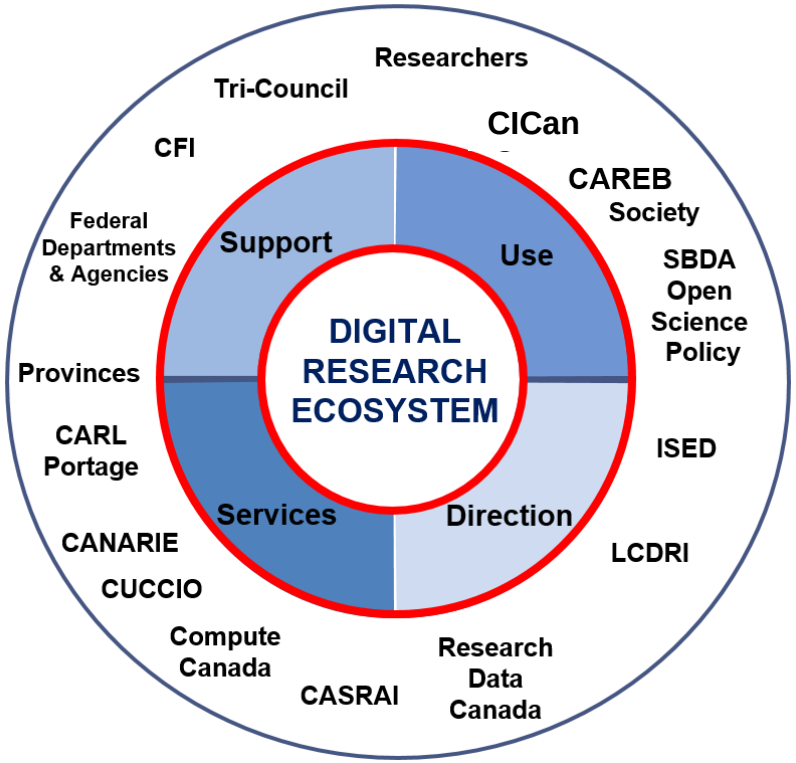
LOCAL

REGIONAL/ PROVINCIAL

NATIONAL



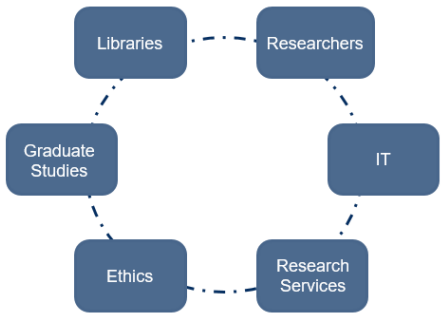


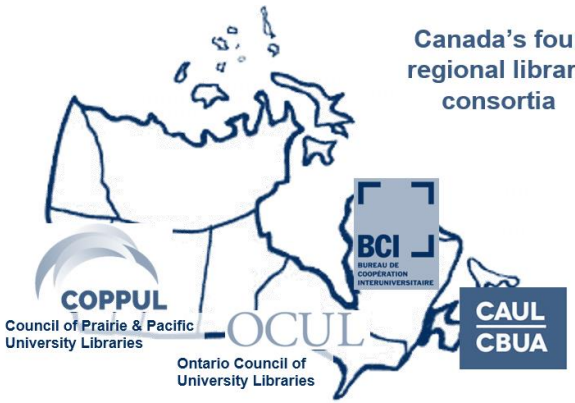


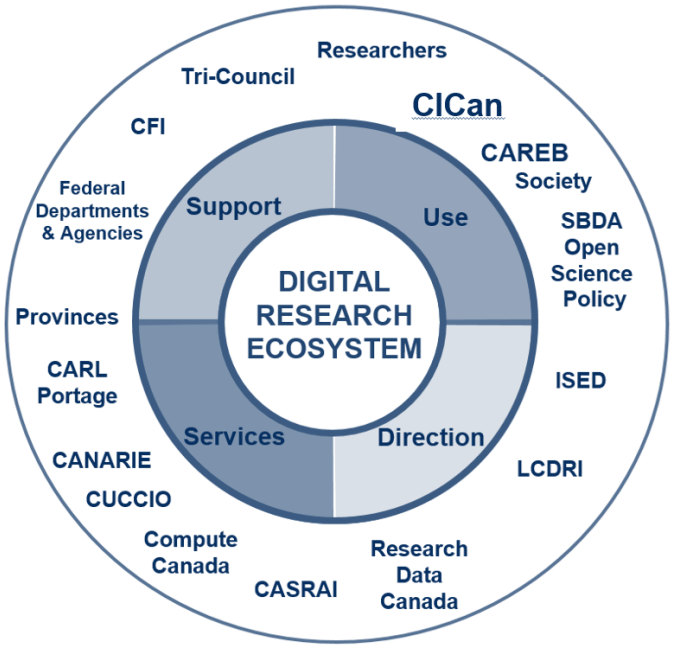
portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

Levels of Stakeholder Engagement

LOCAL | **REGIONAL/ PROVINCIAL** | **NATIONAL**





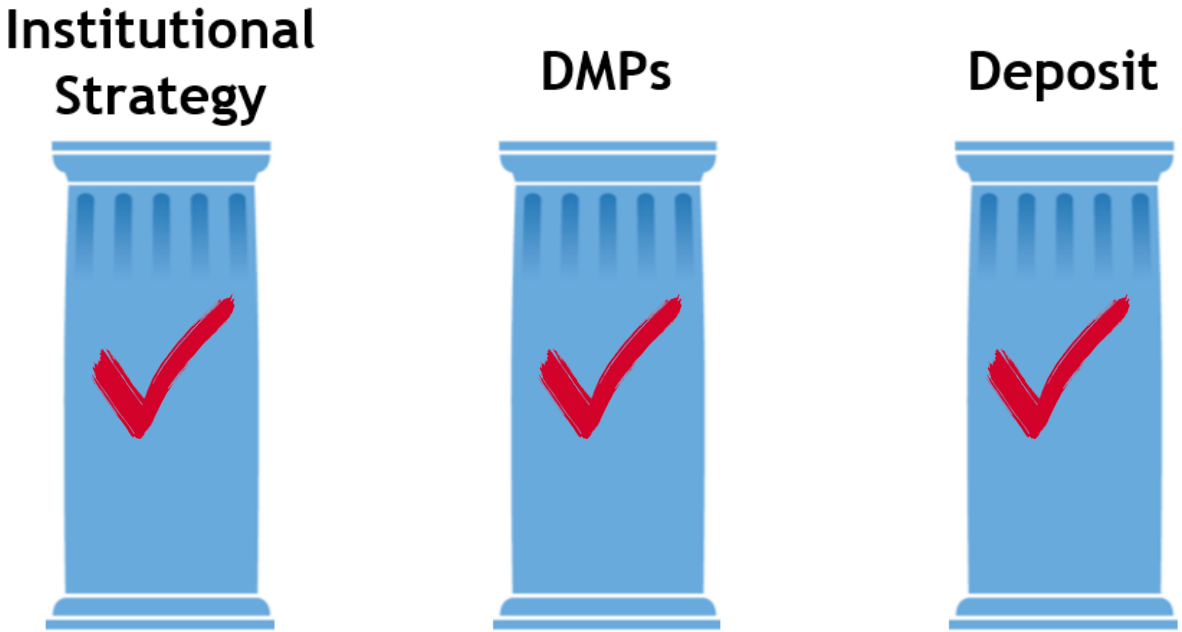


portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

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**

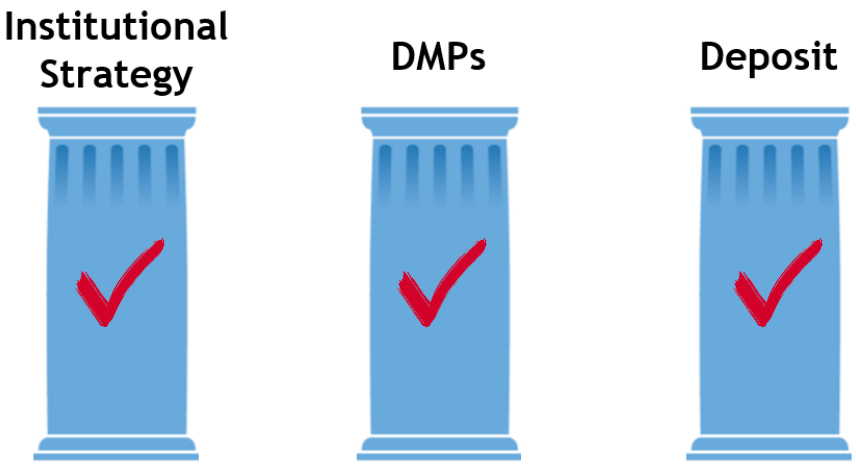
Portage and RDM Infrastructure



Initial focus on the emerging

Tri-Agency RDM Policy

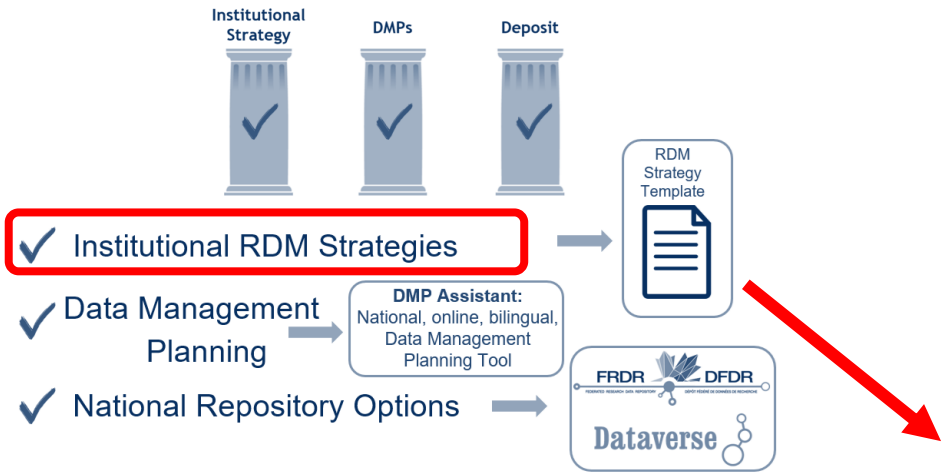
Portage Responses to Tri-Agency Policy



- Institutional RDM Strategies → RDM Strategy Template
- Data Management Planning → **DMP Assistant:** National, online, bilingual, Data Management Planning Tool
- National Repository Options → FRDR (Federated Research Data Repository) / DFDR (Dépôt fédéré de données de recherche) / Dataverse

portage
Services partagés pour les données de recherche
Shared stewardship of research data

Pillar 1: Institutional RDM Strategies



Expert Groups

Data Management Planning

Curation

Data Discovery

Preservation

Training

Research Intelligence

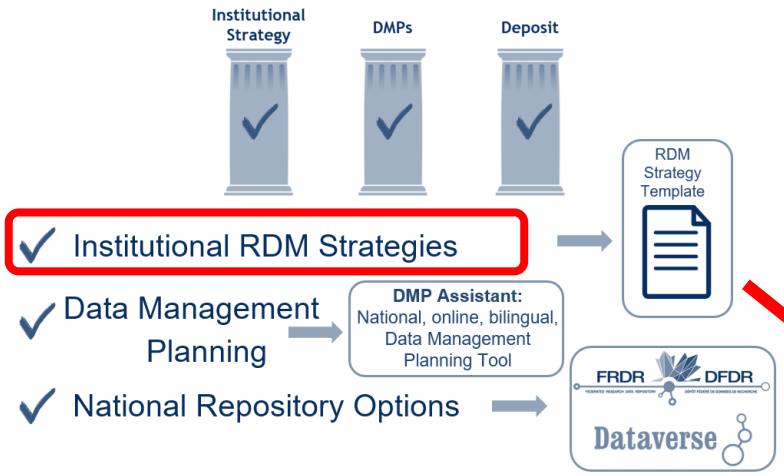
Working Groups

- Dataverse North WG
- Responsible Research Data Management Practices for Sensitive Data
- FRDR Service Model WG
- FRDR Discovery Service WG
- **Institutional RDM Strategies**

RDM Strategy Template

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

Pillar 1: Institutional RDM Strategies



Expert Groups

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

Working Groups

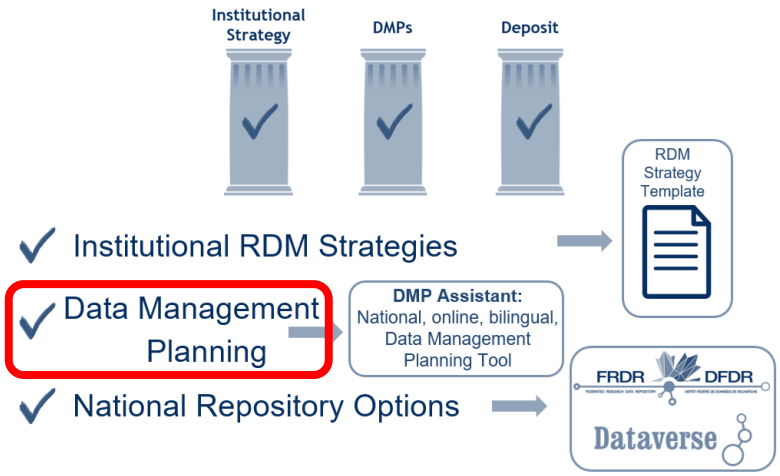
- Dataverse North WG
- Responsible Research Data Management Practices for Sensitive Data
- FRDR Service Model WG
- FRDR Discovery Service WG
- **Institutional RDM Strategies**

Strategy Components | **+ associated guidance**

1. **Raise awareness**: about benefits of best practices and emerging policy requirements.
2. **Assess institutional readiness**: to manage and realize the potential of data holdings.
3. **Formalize RDM practices**: through the development and/or adoption of guidelines, best practices, and policies – setting the tone and underscoring institutional commitment and expectations.
4. **Define a roadmap**: to help with capacity-building over the medium term, to ensure research excellence, efficiency, and transparency.

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

Pillar 2: Data Management Planning



Data Management Planning

Expert Groups

Curation

Data Discovery

Preservation

Training

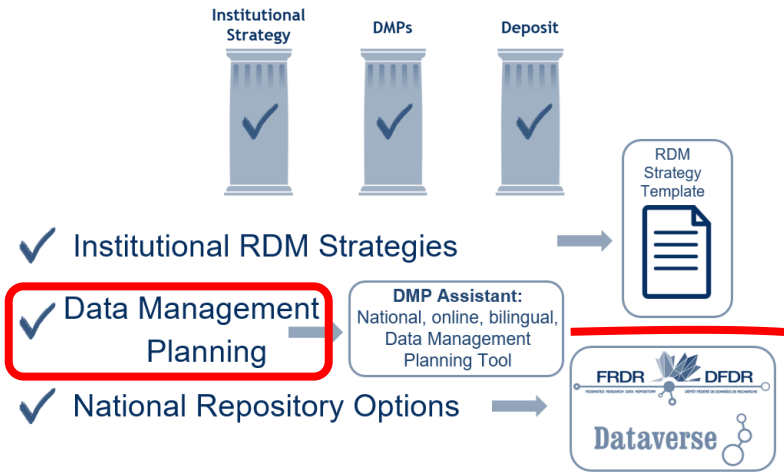
Research Intelligence

Working Groups

- Dataverse North WG
- Responsible Research Data Management Practices for Sensitive Data
- FRDR Service Model WG
- FRDR Discovery Service WG
- Institutional RDM Strategies

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

Pillar 2: Data Management Planning



Working Groups

- Dataverse North WG
- Responsible Research Data Management Practices for Sensitive Data
- FRDR Service Model WG
- FRDR Discovery Service WG
- Institutional RDM Strategies

Data Management Planning

Curation

Expert Groups

Data Discovery

Preservation

Training

Research Intelligence

DMP Assistant
Sign In Sign Up

UNIVERSITY OF ALBERTA

https://assistant.portagenetwork.ca/en

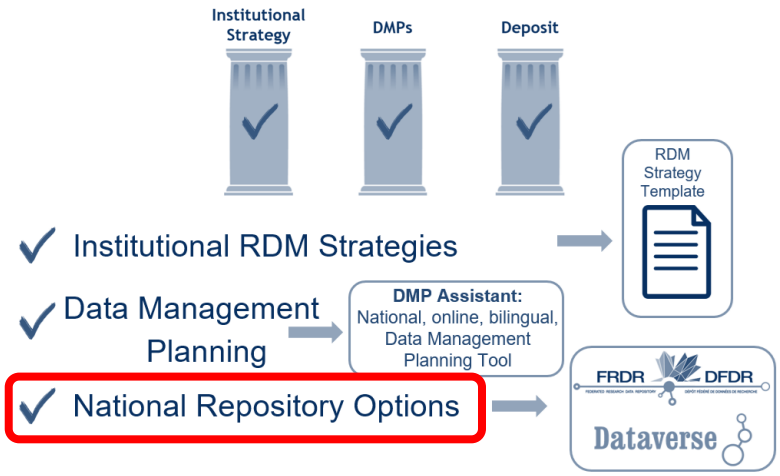
1. **Data collection**
2. **Documentation & metadata**
3. **Backup & storage**
4. **Preservation**
5. **Data sharing & reuse**
6. **Responsibilities & resources**
7. **Ethics & legal compliance**

DMP Assistant 2.0

Coming Soon

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

Pillar 3: Data Deposit into a Repository



Data Management Planning

Expert Groups

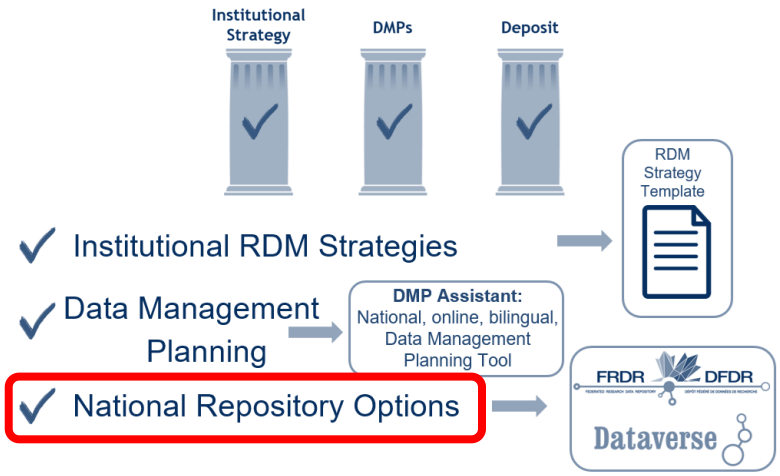
- **Curation**
- **Data Discovery**
- **Preservation**
- **Training**
- **Research Intelligence**

Working Groups

- **Dataverse North WG**
- Responsible Research Data Management Practices for Sensitive Data
- **FRDR Service Model WG**
- **FRDR Discovery Service WG**
- Institutional RDM Strategies

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

Pillar 3: Data Deposit into a Repository



Data Management Planning

Curation

Data Discovery

Expert Groups

Preservation

Training

Research Intelligence

Working Groups

- **Dataverse North WG**
- Responsible Research Data Management Practices for Sensitive Data
- **FRDR Service Model WG**
- **FRDR Discovery Service WG**
- Institutional RDM Strategies

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

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

Canadian Repository Landscape

General Repositories | **Government Repositories** | **Domain-specific Repositories**

Benefits of Data Repositories

From draft Tri-Agency RDM Policy:

“The repository will ensure safe storage, preservation, and curation of the data….Whenever possible, these data, metadata and code should be linked to the publication with a persistent digital identifier.”

Also:

- Promotes best practices for data stewardship

- Researchers not responsible for the ongoing stewardship of their research data in perpetuity

- Increases discoverability and reusability of research data; collect once, reuse many times

- Provides open or controlled access to research data without subscription or licensing fees associated with commercial publishers

Canadian Repository Landscape

General Repositories | **Government Repositories** | **Domain-specific Repositories**

re3data.org

Two National Repository Options

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

Federated Research Data Repository

Dataverse North

National Instance of Dataverse

National Discovery Service

Portage Data Discovery Service Working Group & Compute Canada FRDR Team

compute canada | calcul canada

Portage Dataverse North Working Group and OCUL Scholars Portal Dataverse Team

OCUL Ontario Council of University Libraries

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

Pillar 3: Data Deposit into a Repository

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

Federated Research Data Repository

Main Features:

1. National Repository Option

'Big Data' upload/download

Maintains file hierarchies

Scalable

2. National Discovery Platform

3. Preservation Processing capacity

Currently in Limited Production

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

MERIDIAN

"Big data solutions for the marine acoustic community"
Dalhousie University

Global Institute for Water Security
Water is life. And water security - both quality and quantity is one of the most critical issues facing our planet.

the water INSTITUTE University of Waterloo
The Water Institute
U. Of Waterloo

MOUNTAIN LEGACY PROJECT

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

compute canada | calcul canada

https://portagenetwork.ca/frdr-dfdr

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

Pillar 3: Data Deposit into a Repository

Limited Production by the numbers:

- Working with 18 Limited Production partners
- 52 datasets published (4.2 Tb)
- 32 datasets in progress (2.7 Tb)

Google Analytics for https://www.frdr.ca/ (May-Oct 2018)

- Users: 2,937/Avg. per month: 490
- Sessions: 5,001/Avg. per month: 834
- Pageviews: 18,057/Avg. per month: 3,010

https://portagenetwork.ca/frdr-dfdr

Pillar 3: Data Deposit into a Repository

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

Federated Research Data Repository

Main Features:

1. National Repository Option
   - ‘Big Data’ upload/download
   - Maintains file hierarchies
   - Scalable
2. National Discovery Platform
3. Preservation Processing capacity

Currently in Limited Production

Dataverse

Main Features:

- Mature platform
- Widely adopted
- User-friendly
- Brandable

National Repository Option

OCUL Ontario Council of University Libraries

Scholars Portal
A Service of the Ontario Council of University Libraries

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

https://portagenetwork.ca/frdr-dfdr

Pillar 3: Data Deposit into a Repository

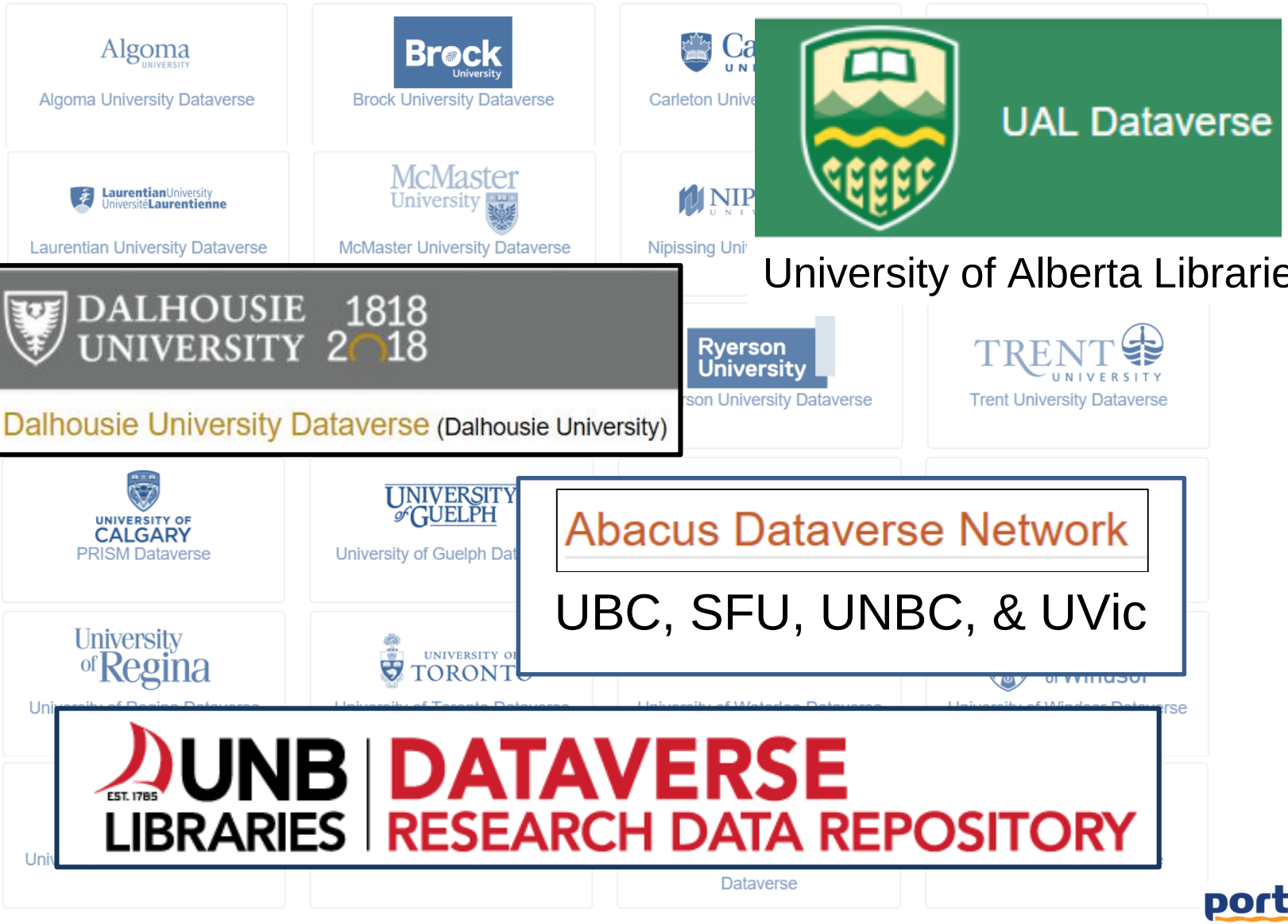
Dataverses at OCUL Scholars Portal:

- Algoma University Dataverse
- Brock University Dataverse
- Carleton University Dataverse
- Lakehead University Dataverse
- Laurentian University Dataverse
- McMaster University Dataverse
- Nipissing University Dataverse
- OCAD University Dataverse
- Queen's University Dataverse
- Royal Military College Dataverse
- Ryerson University Dataverse
- Trent University Dataverse
- PRISM Dataverse
- University of Guelph Dataverse
- UOIT Dataverse
- University of Ottawa Dataverse
- University of Regina Dataverse
- University of Toronto Dataverse
- University of Waterloo Dataverse
- University of Windsor Dataverse
- University of Windsor Dataverse
- Western University Dataverse
- Wilfrid Laurier University Dataverse
- York University Dataverse

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

Pillar 3: Data Deposit into a Repository

… Dalhousie, Abacus, & UNB



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

Pillar 3: Data Deposit into a Repository

Who to contact?

Interested in joining Dataverse North?

Contact:

dataverse@scholarsportal.info

for more information!

UNB LIBRARIES | DATAVERSE RESEARCH DATA REPOSITORY

University of Windsor Dataverse

Western University Dataverse

Wilfrid Laurier University Dataverse

York University Dataverse

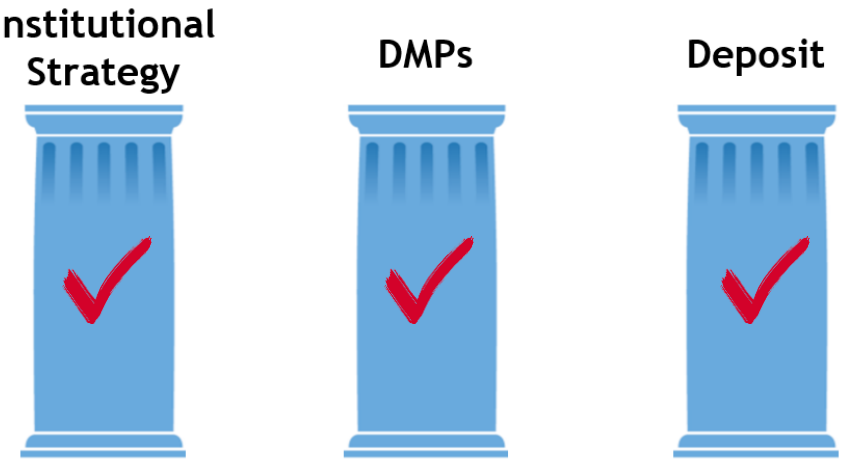
portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

Portage RDM Support for Institutions & Researchers

Tri-Agencies

&

Portage



Portage

Portage

RDM Strategy Template

DMP Assistant

Sign In | Sign Up

National, online, bilingual, Data Management Planning Tool

FRDR — FEDERATED RESEARCH DATA REPOSITORY

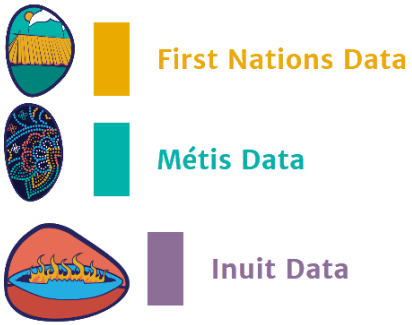
DFDR — DÉPÔT FÉDÉRÉ DE DONNÉES DE RECHERCHE

Dataverse

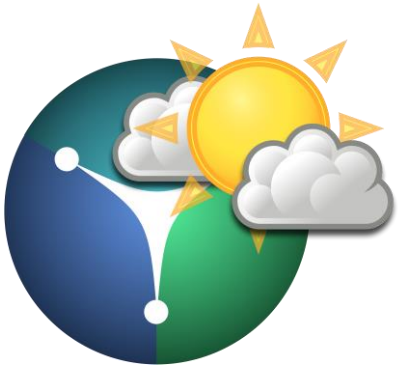
portage

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

Circling back to another Portage initiative…



Sensitive Data



General
Data
Protection
Regulation

https://thehackernews.com/2018/04/facebook-data-privacy.html
https://www.opengrid.com/content/2013/01/cim-for-environmental-data/cim-for-weather-logo/
http://www.prweb.com/releases/2014/07/prweb12048854.htm
http://www.startribune.com/the-trans-pacific-partnership-and-intellectual-property-an-opportunity-lost/416474353/
https://www.ices.on.ca/About-ICES/Collaborations-and-Partnerships/Indigenous-Portfolio

Portage ‘Sensitive Data’ Working Group

- ✓ Institutional RDM Strategies WG
- ✓ DMP Assistant
- ✓ National Repository Options

National, online, bilingual, Data Management Planning Tool

RDM Strategy Template

FRDR DFDR

Dataverse

Expert Groups

Data Management Planning

Curation

Data Discovery

Preservation

Training

Research Intelligence

Working Groups

- Dataverse North WG
- **Responsible Research Data Management Practices for Sensitive Data**
- FRDR Service Model WG
- FRDR Discovery Service WG
- Institutional RDM Strategies

https://en.wikipedia.org/wiki/File:Johnny-automatic-scales-of-justice.svg

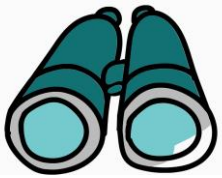
portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

Four priority outcomes identified…

Deposit-friendly text for ethics applications & informed consent

Data access agreements

Training materials



Environmental scan of best practices for Indigenous data

https://www.operational-excellence-consulting.com/opex-academy
https://eyetube.net/podcasts/informed-consent/
https://www.freepik.com/free-icon/text-document_735526.htm

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**

Vision for a Federated RDM Architecture

Outline

- Federal RDM policy developments
- Further context for RDM in Canada
- Goals and outputs of Portage
- **Developing a shared vision for RDM in Canada**

Federated RDM Architecture

Active Storage

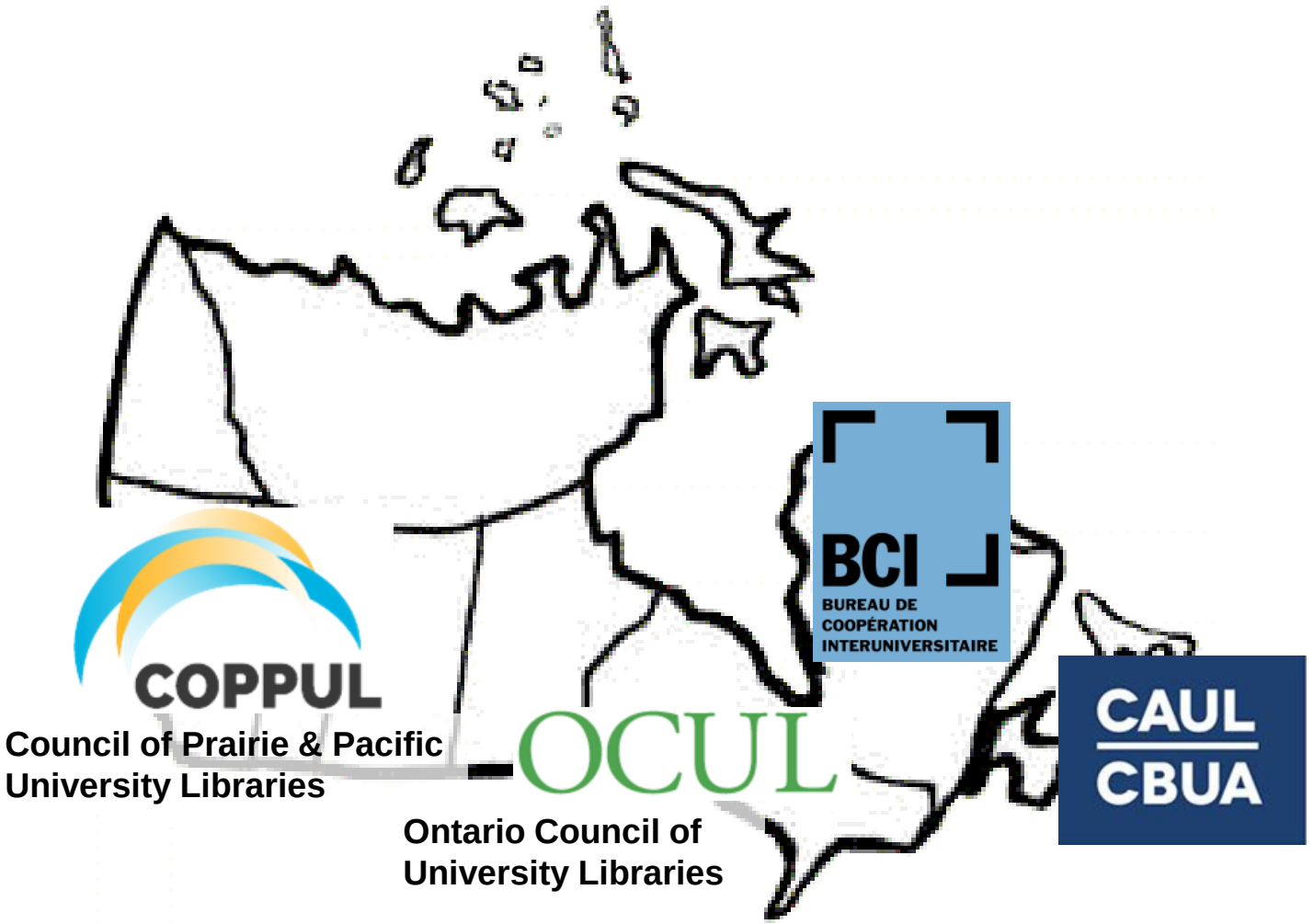
Repository Storage

Archival (Preservation) Storage

Data Management Planning

Data Curation

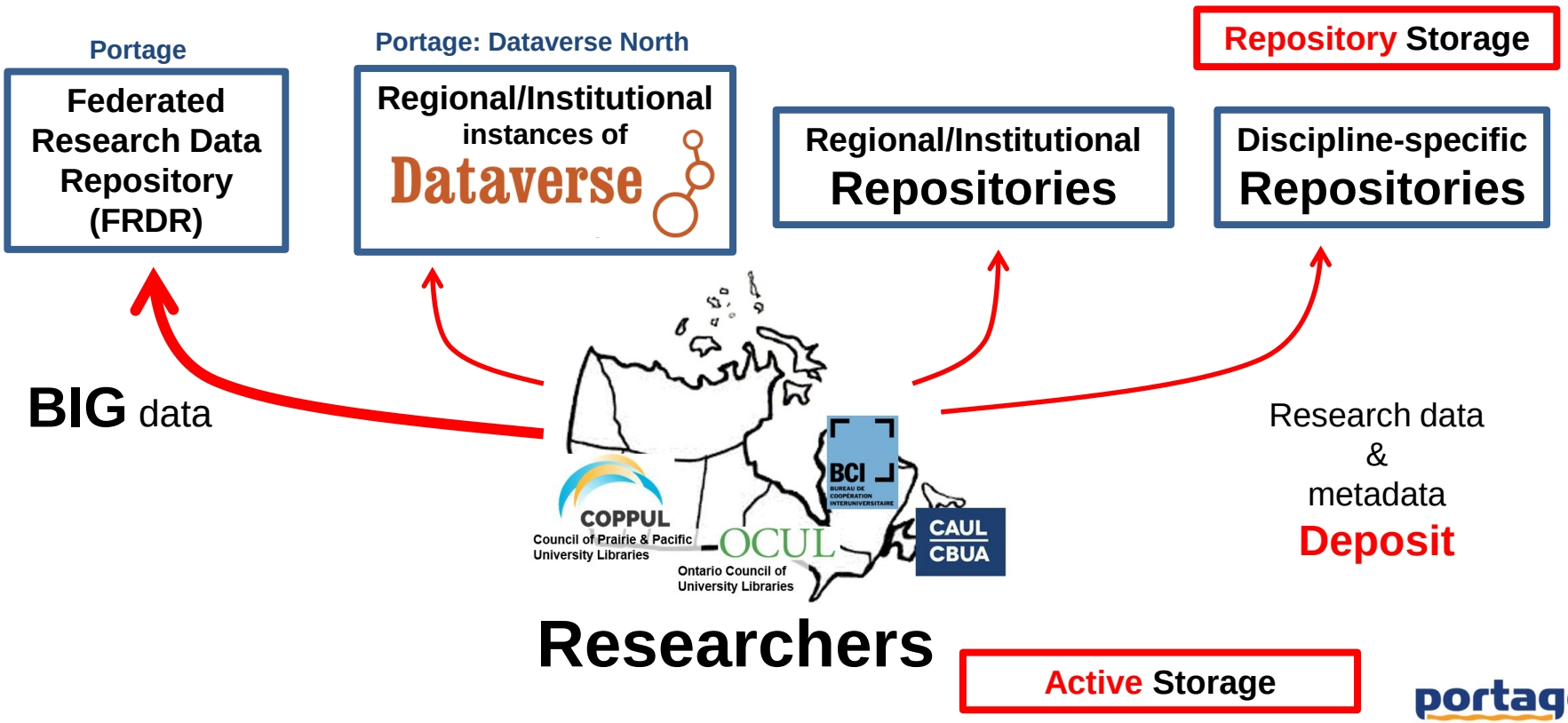
Federated RDM Architecture



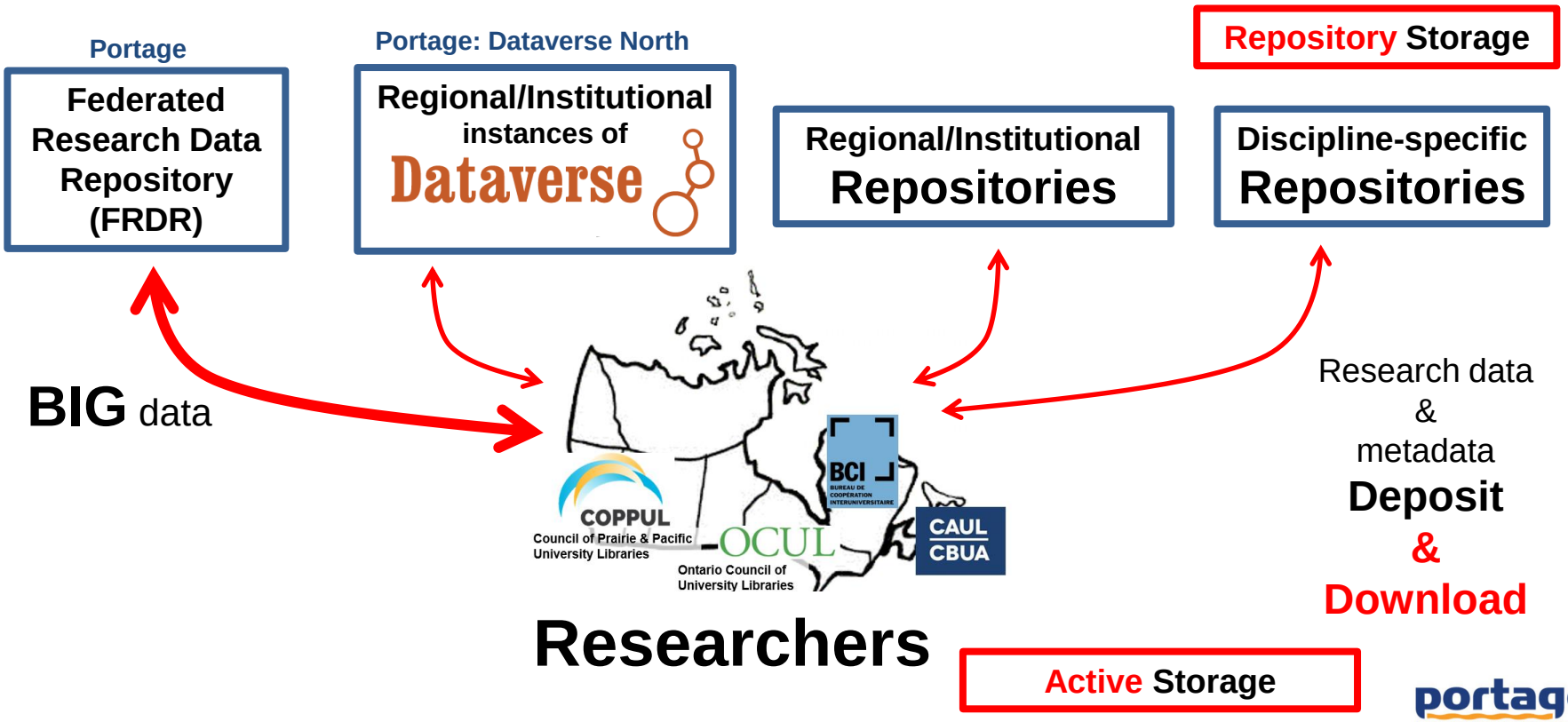
Researchers

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

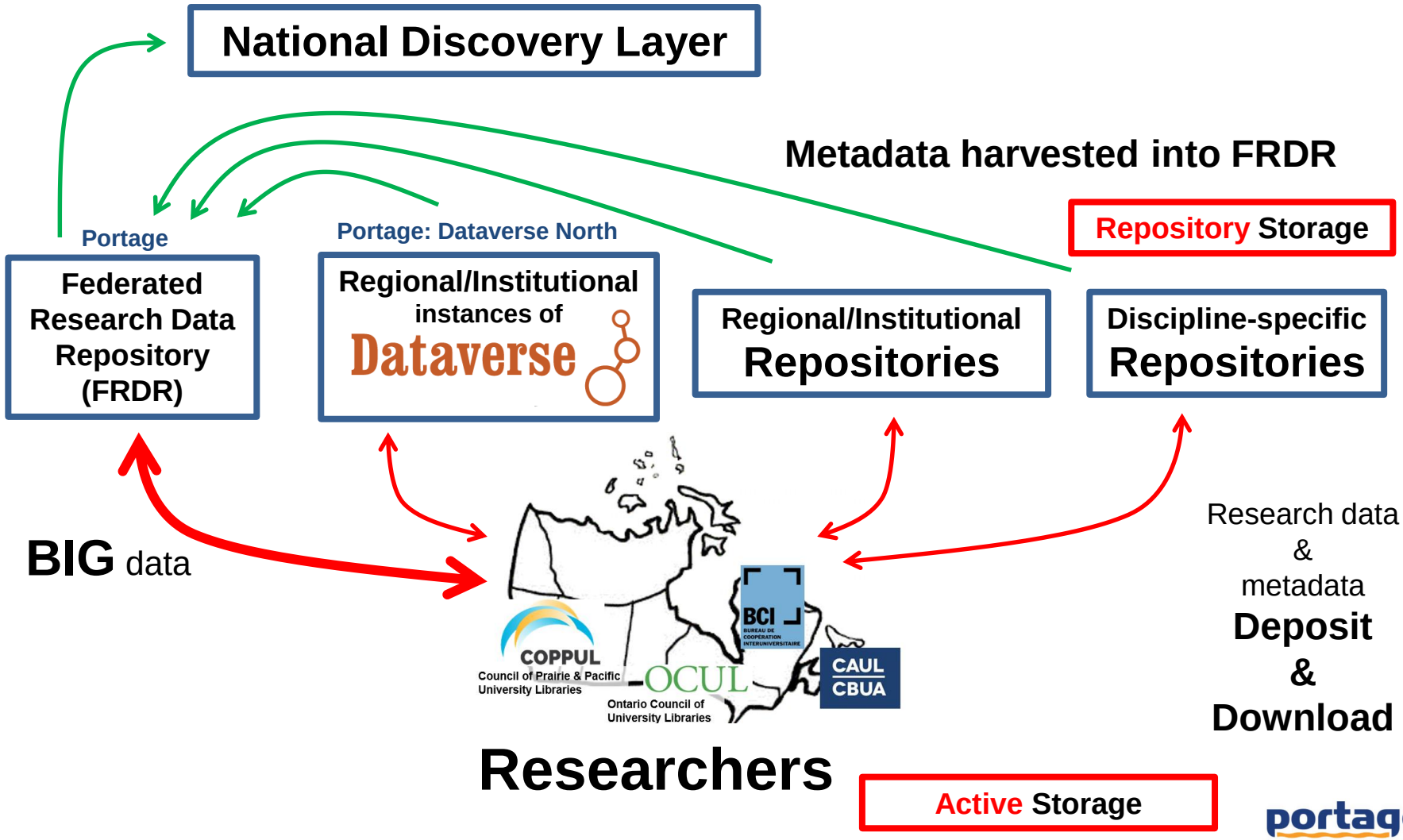
Federated RDM Architecture



Federated RDM Architecture



Federated RDM Architecture



Discovery: Collaborating repositories

Metadata harvested from 40 repositories indexing over 90,000 datasets

All searchable via https://www.frdr.ca

National Discovery Layer

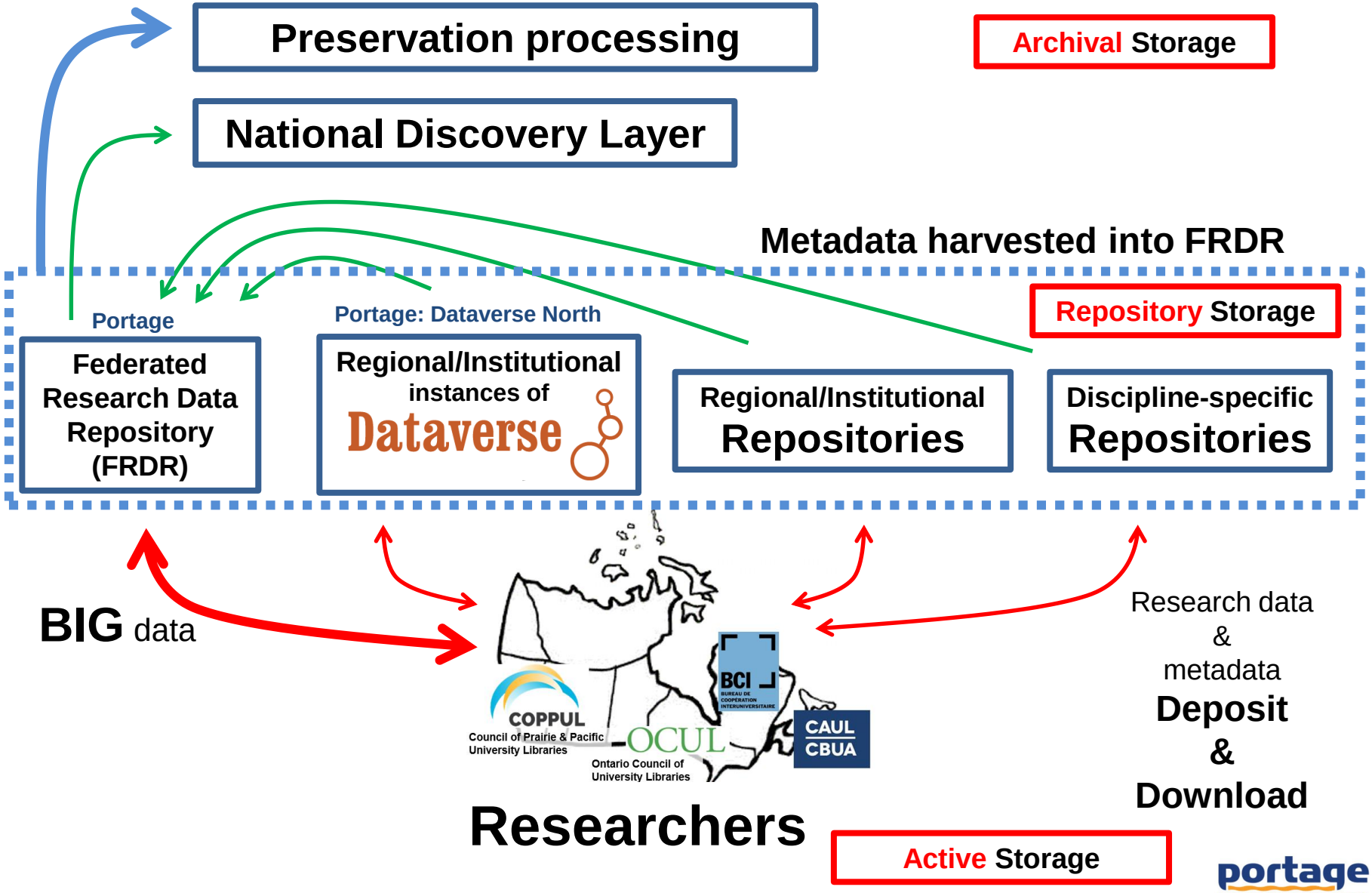
- FRDR's harvester indexes data repositories from across Canada to make research data held in many repositories discoverable from a single platform
- Bilingual user interface for search and discovery
- Indexing 40+ repositories; 90,000 metadata records

- **Metadata ≠ Data**

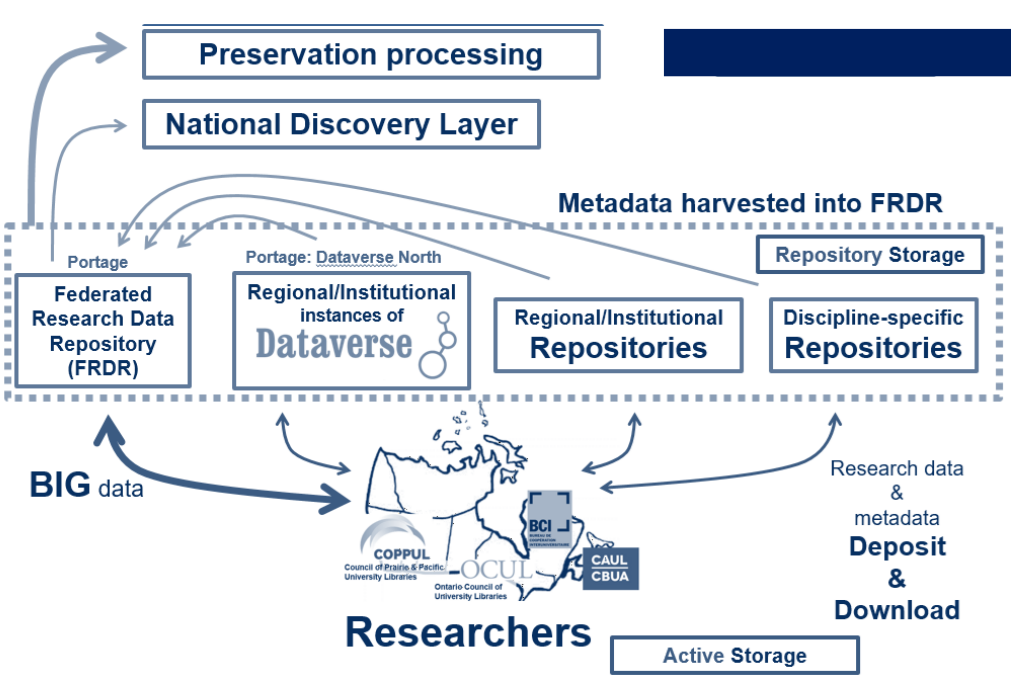
Goals:

- **Improve discovery** of research data/metadata across Canada
- **Drive traffic** to existing repository sites
- **Break down repository siloes**

Federated RDM Architecture



Federated RDM Architecture



Archival Storage

Preservation Processing via Archivematica

Convert to 'future-friendly' formats (e.g. docx → PDF/A)

&

Generation of Archival Information Packages (AIPs)

Storage in a Network of Preservation Service Providers

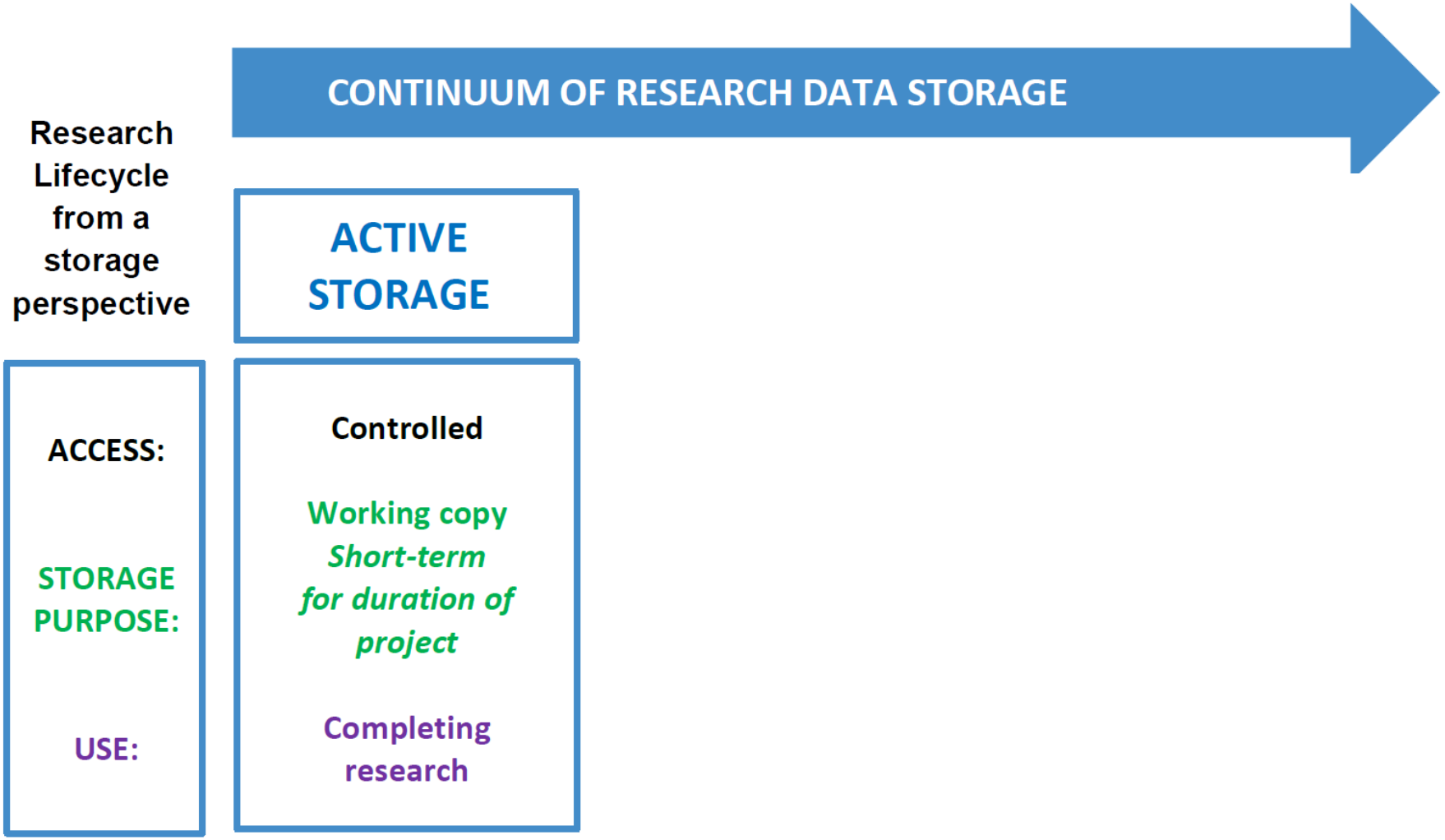
COPPUL | Council of Prairie and Pacific University Libraries

WestVault

ONTARIO LIBRARY RESEARCH CLOUD

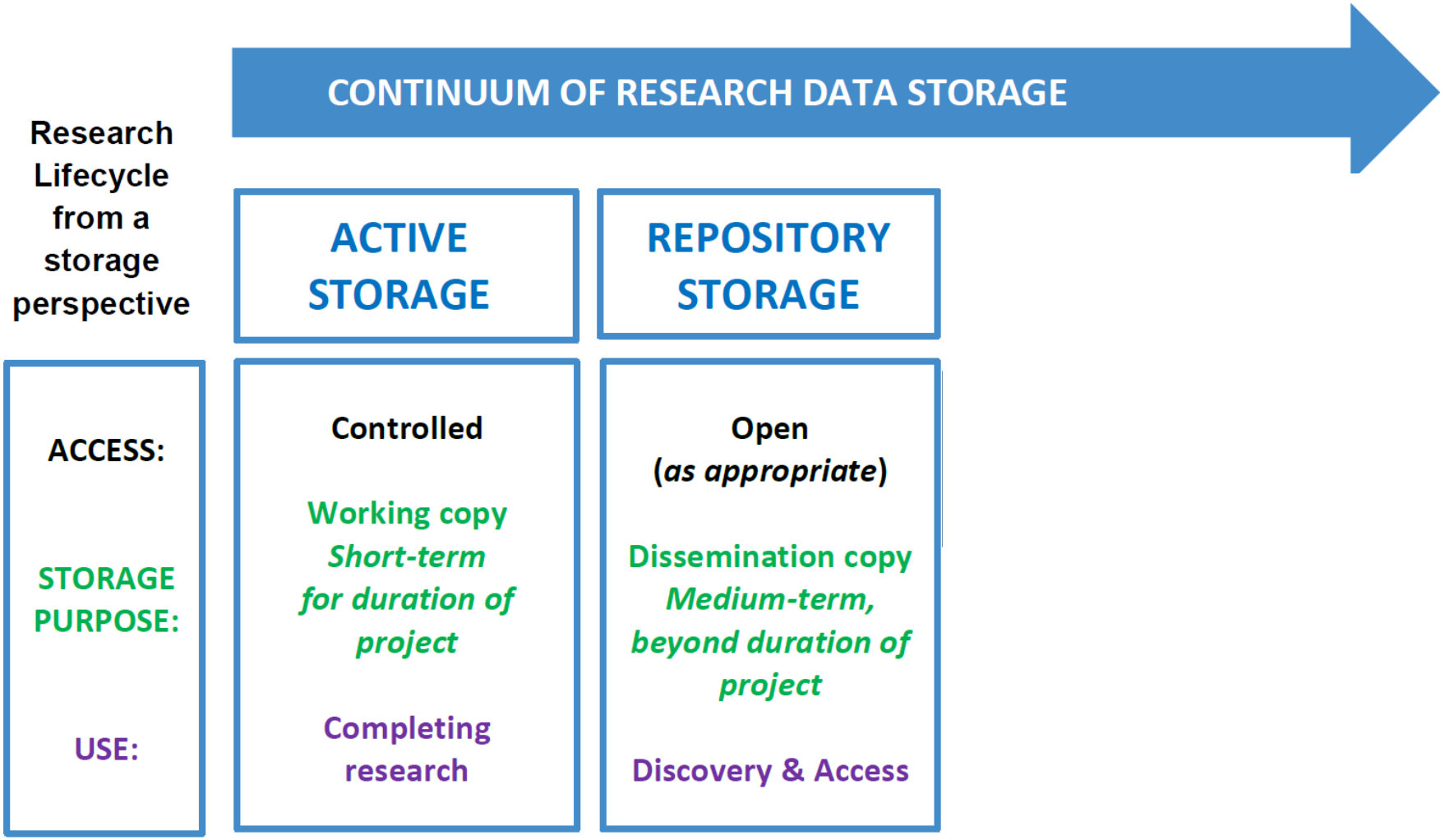
portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

Data Storage Continuum



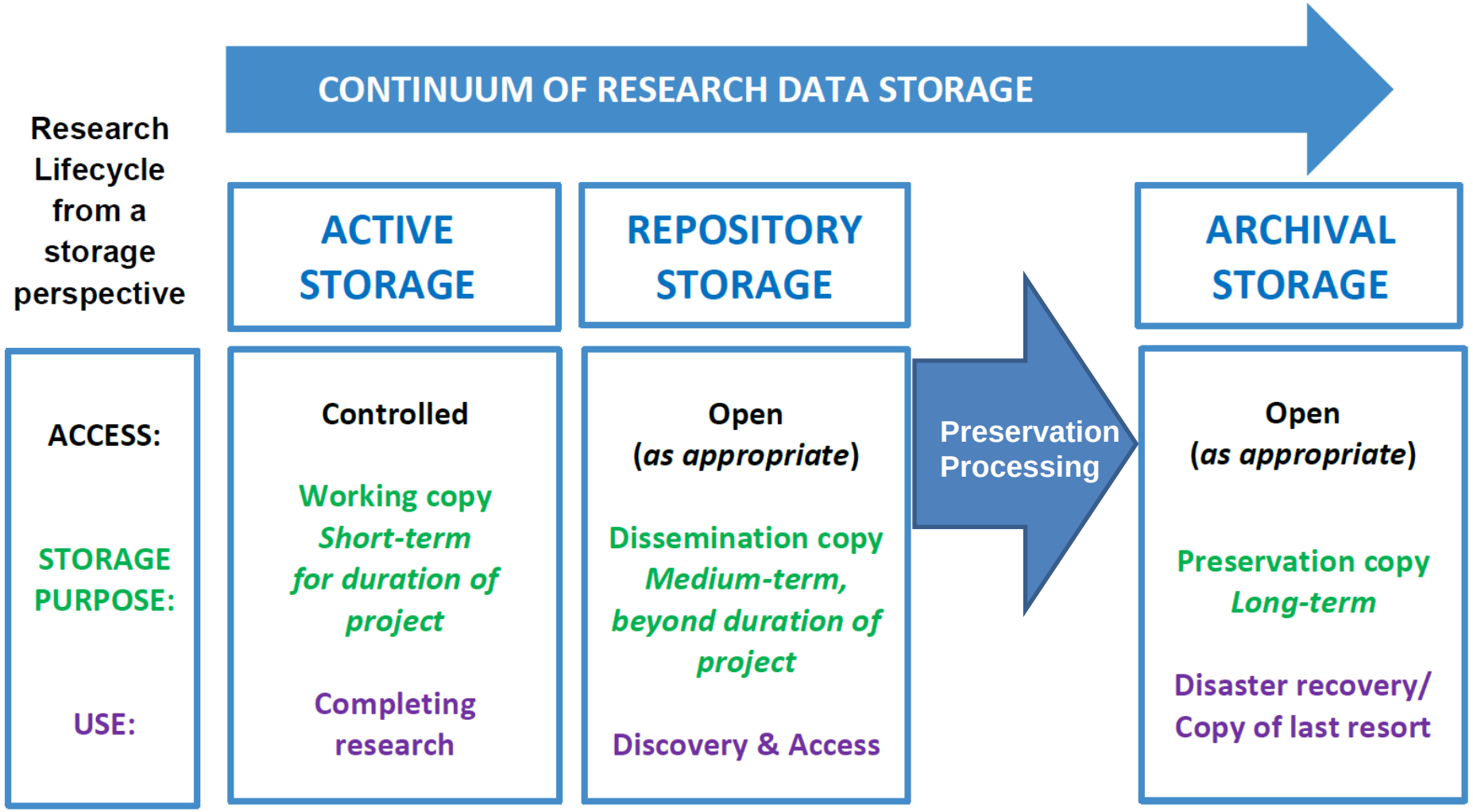
portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

Data Storage Continuum



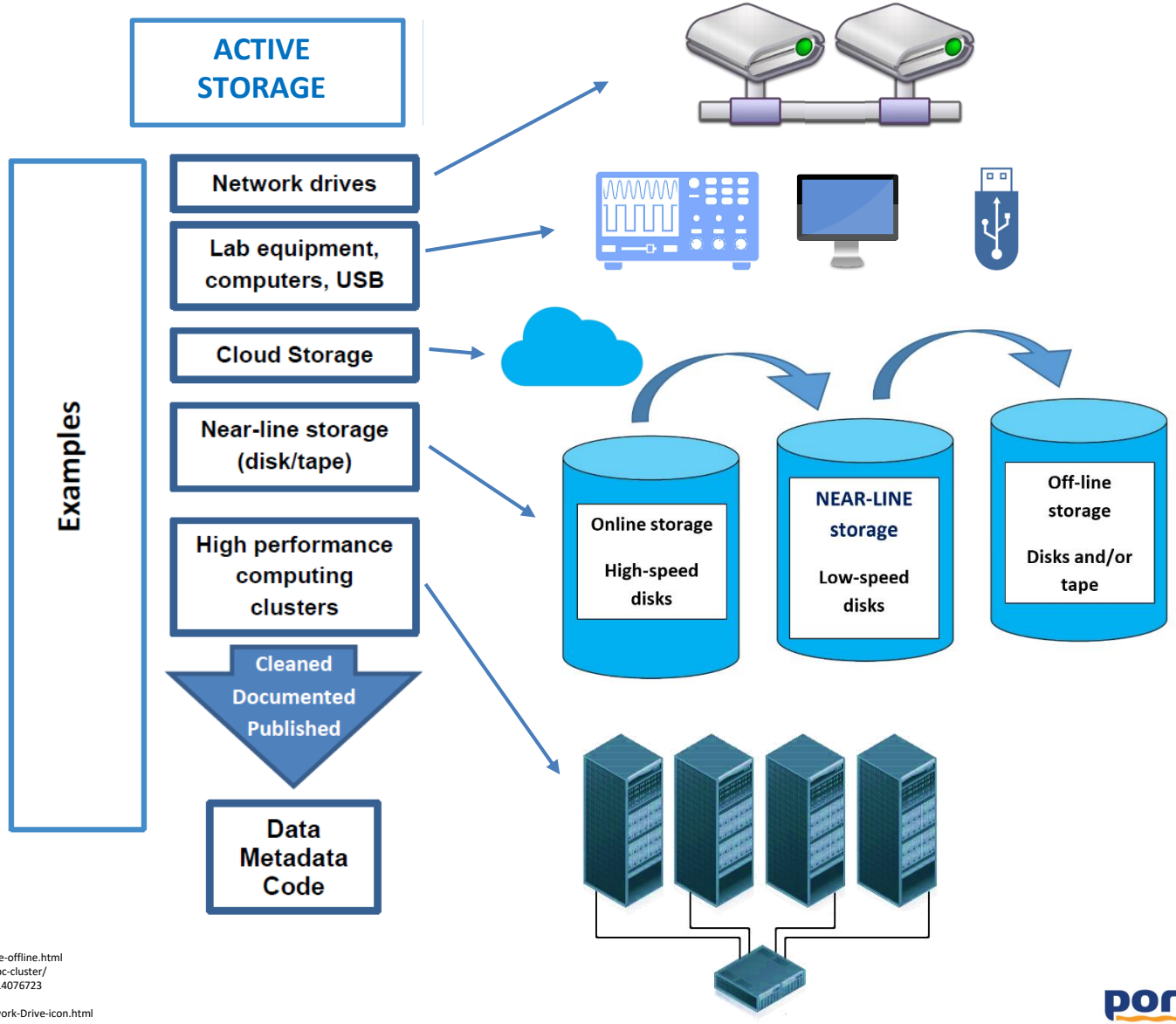
portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

Data Storage Continuum



Examples of each type of storage…

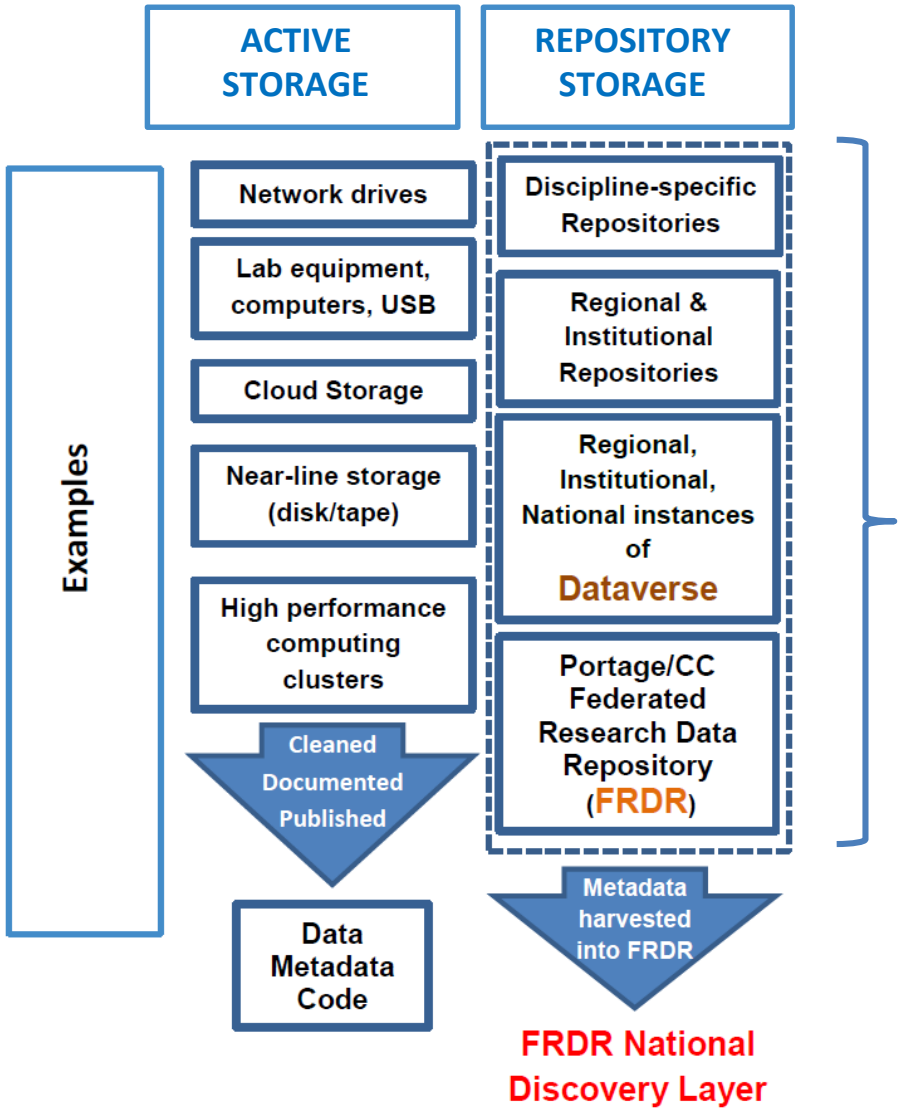
ACTIVE STORAGE supporting completion of research



http://weeratapwutti.blogspot.com/2014/12/online-storage-nearline-storage-offline.html
https://www.happyware.com/it/it-solutions/cluster-solutions/clustertyps/hpc-cluster/
https://www.vectorstock.com/royalty-free-vector/oscilloscope-icon-vector-14076723
https://101clipart.com/clipart-computer-monitor/
http://www.iconarchive.com/show/refresh-cl-icons-by-tpdkdesign.net/Network-Drive-icon.html

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

Examples of each type of storage…

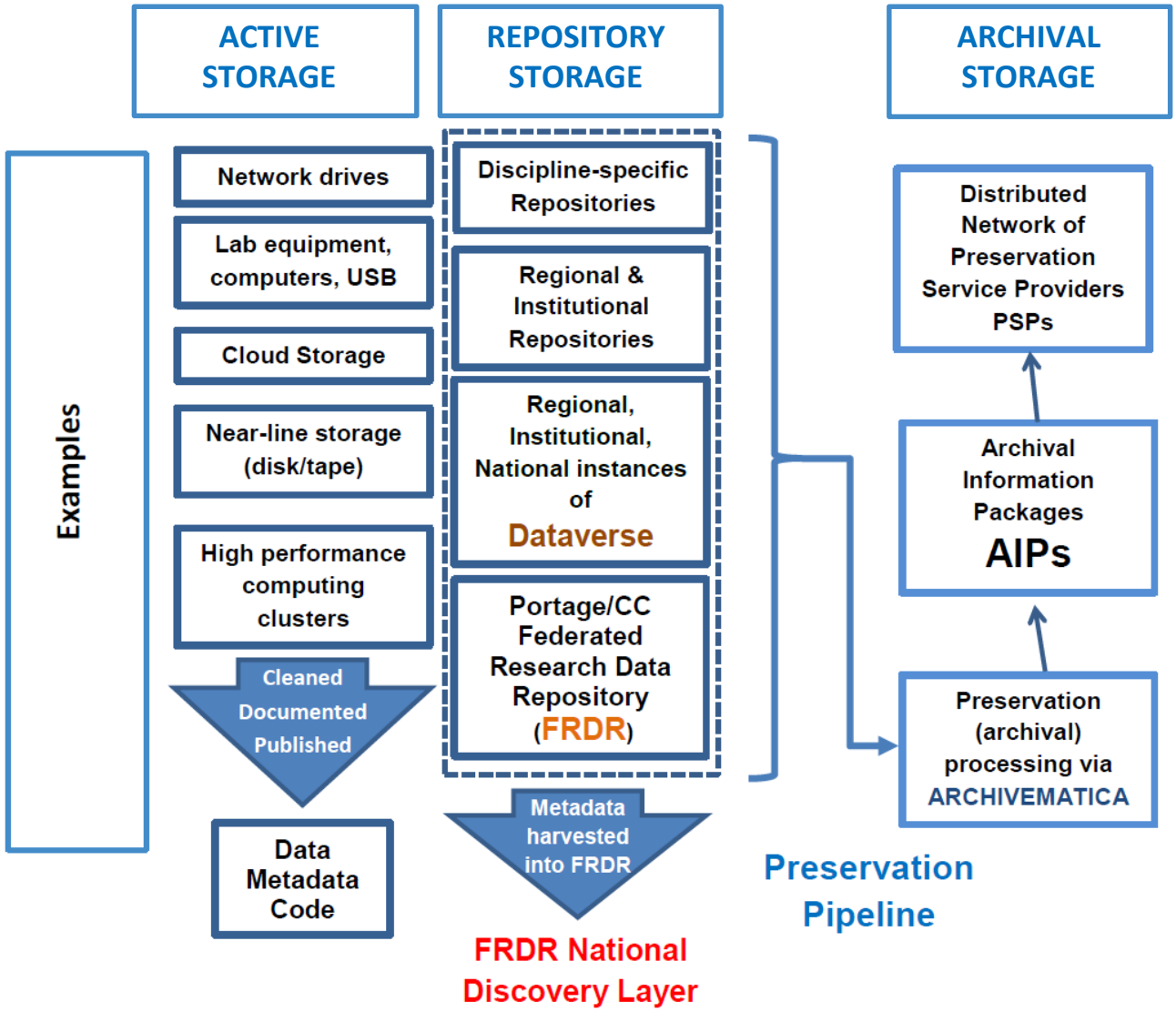


REPOSITORY STORAGE supporting discovery & access

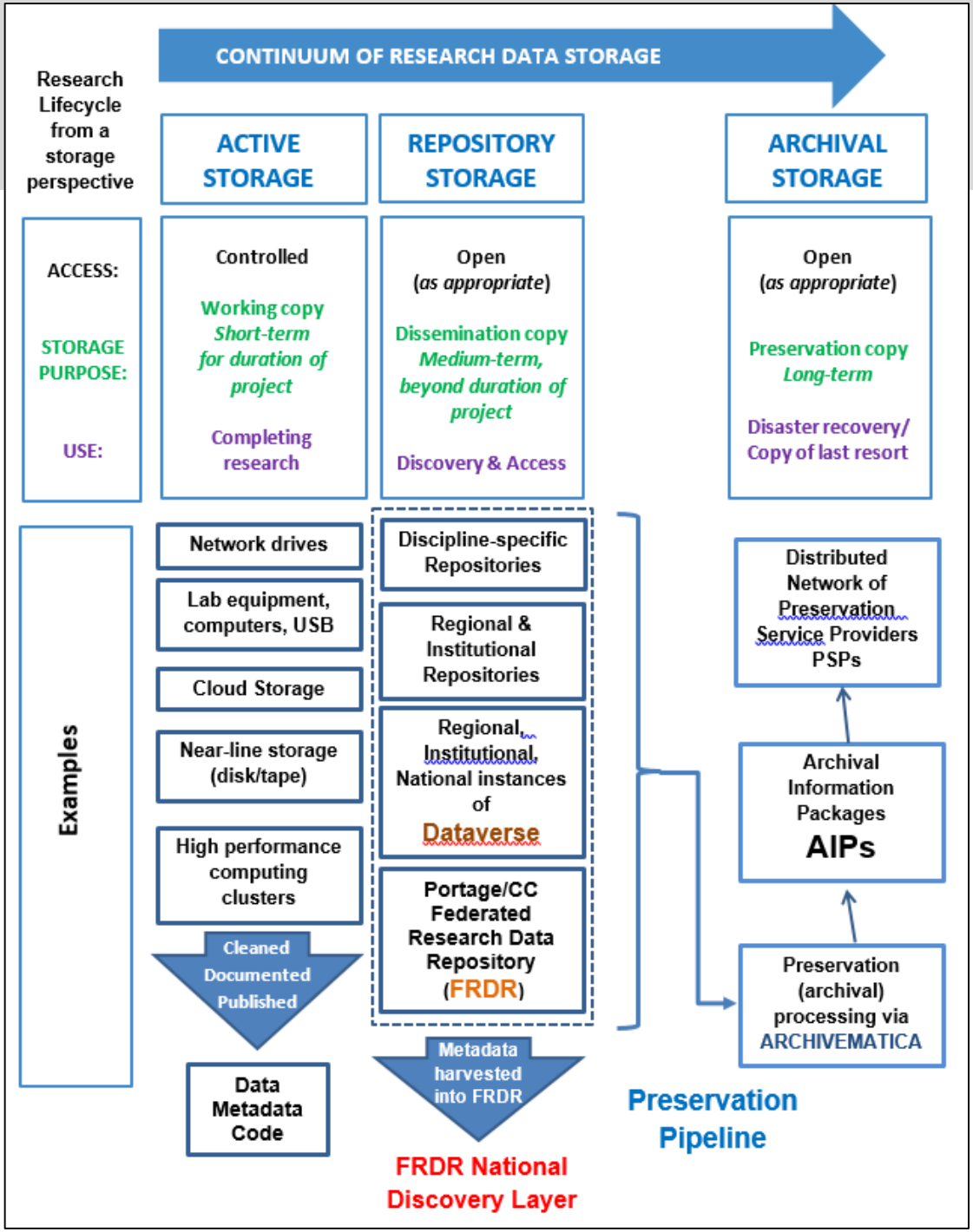
portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

Examples of each type of storage…

ARCHIVAL STORAGE supporting long-term preservation



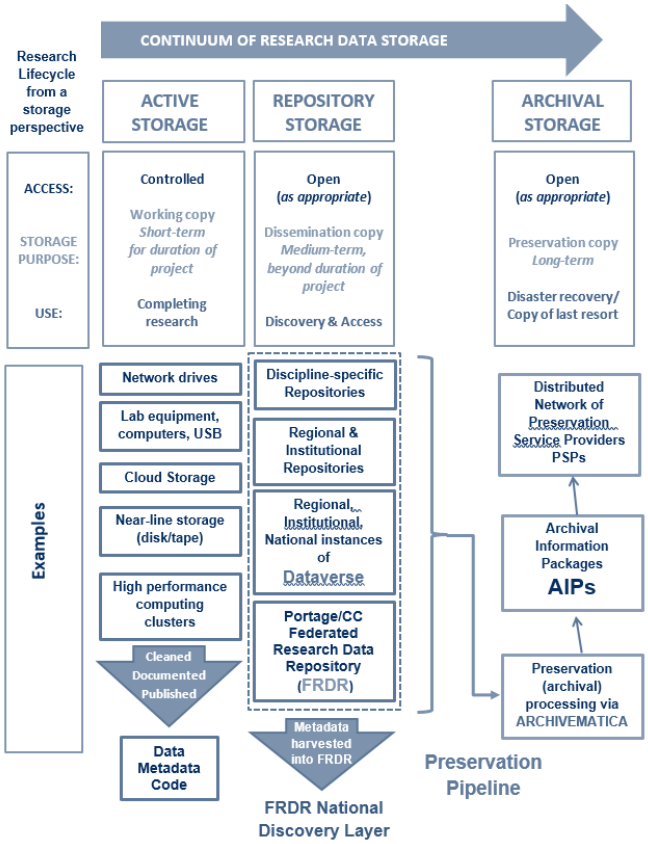
Data Storage Continuum



A useful way of framing the storage picture for RDM

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

But other activities happen along the way…



And the whole can be described as

GOOD DATA STEWARDSHIP

DATA MANAGEMENT PLANNING

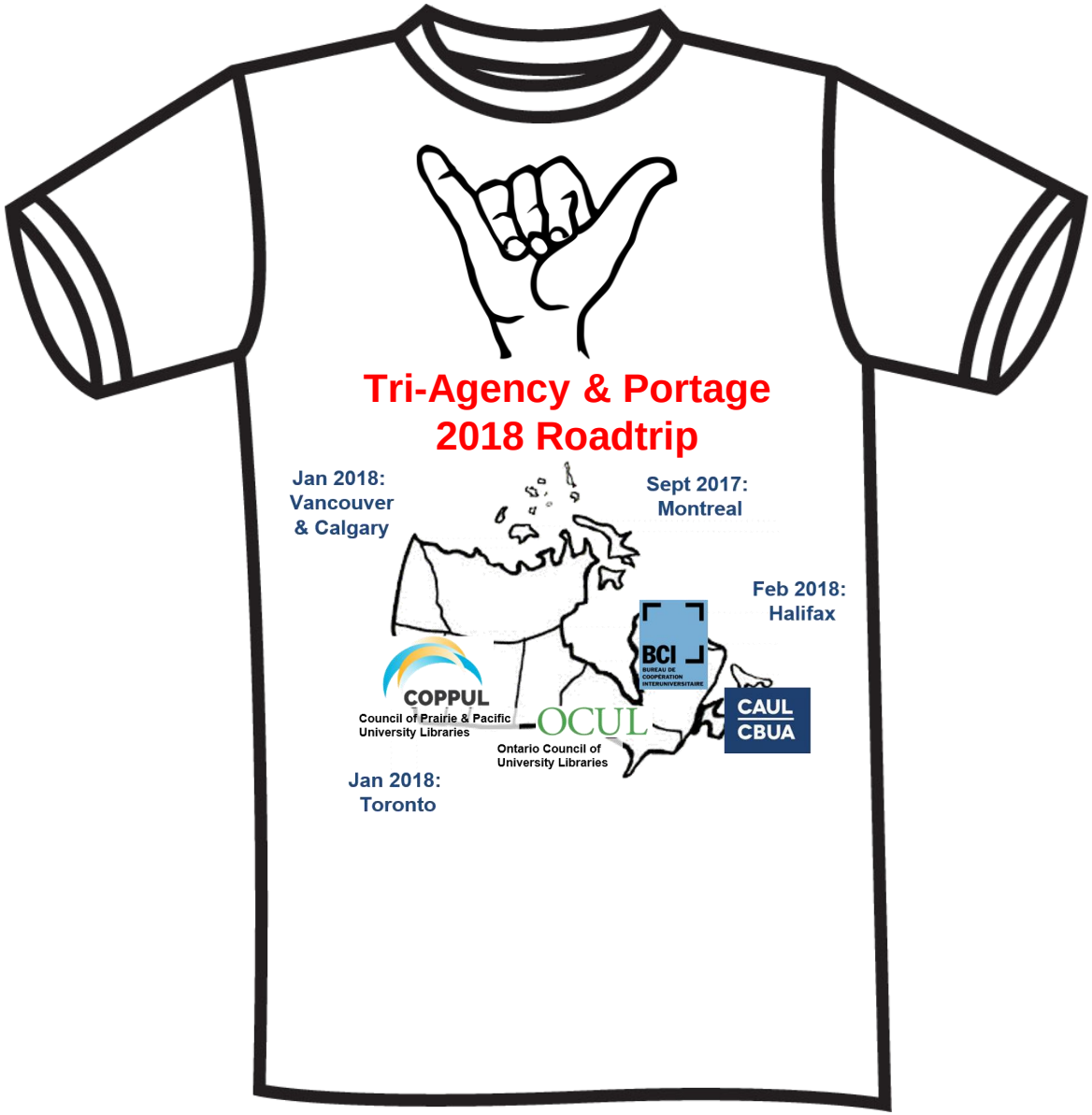
DATA CURATION

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

Staying in touch with RDM stakeholders…

Portage Outreach

Portage Outreach with the Tri-Agencies



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

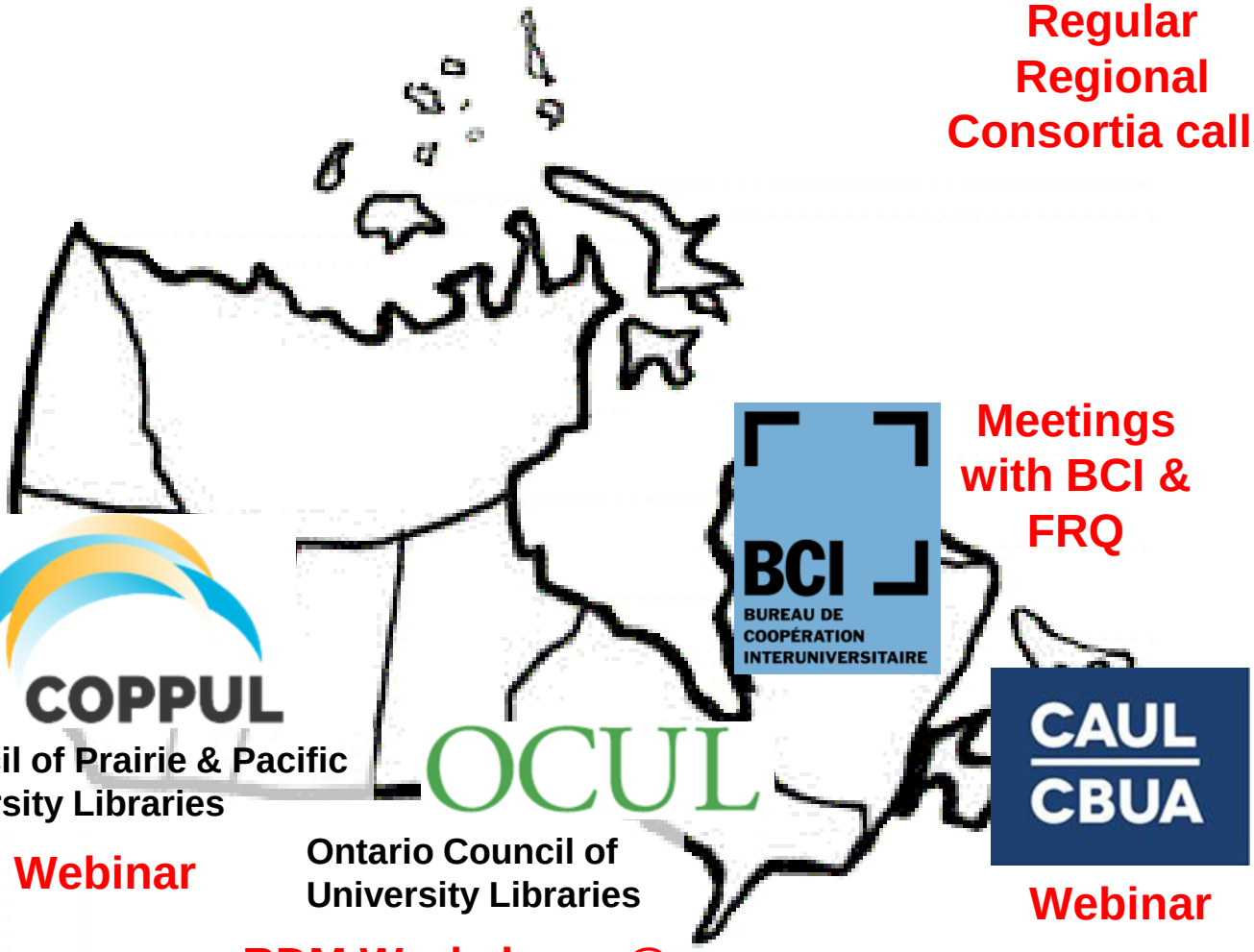
Portage Outreach: *national and diverse*

RDC DRC
Research Data Canada – Données de recherche Canada

Regular updates & calls
Steering Committee
Advisory Committee
Infrastructure Committee

CANARIE RDM Funding Call

Regular Regional Consortia calls



Meetings with BCI & FRQ

BCI
BUREAU DE COOPÉRATION INTERUNIVERSITAIRE

COPPUL
Council of Prairie & Pacific University Libraries

Webinar

OCUL
Ontario Council of University Libraries

CAUL CBUA

Webinar

CICan Applied Research Symposium

RDM Workshops @ WLU and Lakehead

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

Successful CANARIE RDM Proposals

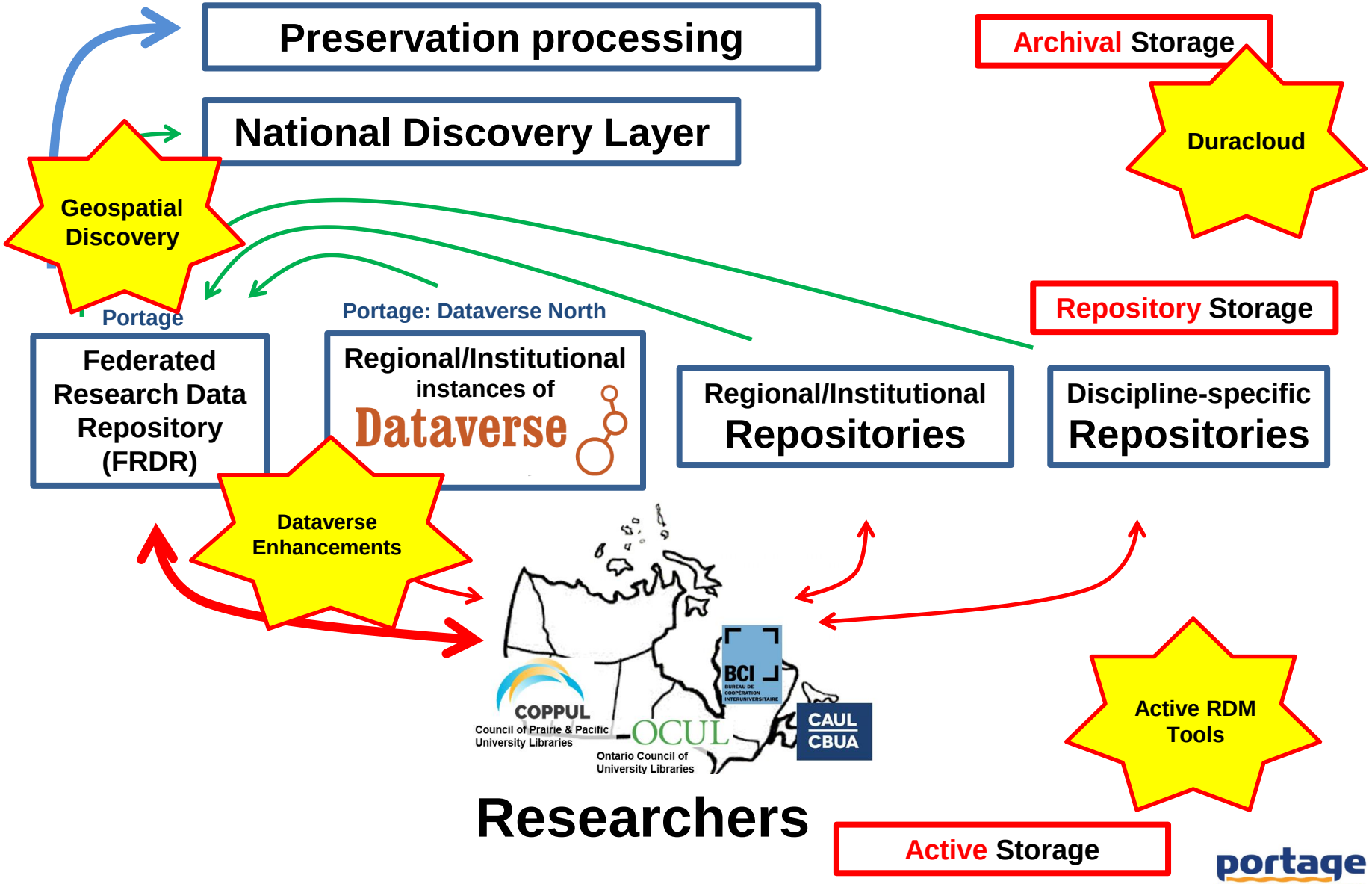
- A **National Geospatial Discovery Layer** to improve the discoverability of geospatial metadata harvested by FRDR and create connections between Dataverse and a GeoServer to enhance Dataverse's ability to visualize geospatial information and generate geocoded metadata for subsequent harvesting by FRDR. Proposal is led by UBC with support from Compute Canada and Scholar's Portal.

- **Improved integration with 'Active' RDM tools** through metadata aggregation and enhancement plug ins/APIs (for example, HubZero and Open Science Framework). Proposal to be led by the Compute Canada RDM team at USask (the same team that developed FRDR).

- **Dataverse enhancements for national repository deployment** - enhance features in Dataverse to make it a better fit for use as a national repository option, including integration with Shibboleth, with the OLRC, and simplified workflows for uploading data, especially large datasets (i.e., integration with Globus File Transfer and Compute Canada storage infrastructure). Proposal led by Scholar's Portal.

- **Data Preservation:** Enhance DuraCloud to be deployed outside the AWS system and add integration for multiple Canadian research data repositories and storage platforms. Led by U of Toronto.

Federated RDM Architecture



Portage Outreach: *lots more on the go!*

Conferences, Symposia, AGMs

- *G7 Research Summit*
- *Data Curation Workshop*
- *CNC CODATA AGM*
- *IASSIST & RDA*
- *Colleges – Applied Research Symposium*
- *College VPR Summit*
- *Statistics Canada @ 100 panel*
- *NDSF Summit*
- *OLA Super Conference*
- *Open Science Conference*
- *IASSIST*

Regular and Targeted Meetings

- *Portage-ISED meetings*
- *Portage-Compute Canada meetings*
- *Portage-RDC meetings*
  - *DM 'Dream Team' meeting*
  - *regular monthly calls with RDC*
- *Portage-Regional Consortia calls*
- *Portage-Scholars Portal calls*

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

Portage: Partnerships & Progress

compute canada | calcul canada

OCUL Ontario Council of University Libraries

Scholars Portal
A Service of the Ontario Council of University Libraries

UNIVERSITY OF ALBERTA

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

Dataverse

DMP Assistant
Sign In Sign Up

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

RDC DRC
Research Data Canada – Données de recherche Canada

CANARIE RDM Funding

- Geospatial Discovery
- Duracloud
- Dataverse Enhancements
- Active RDM Tools

CIHR IRSC
Canadian Institutes of Health Research | Instituts de recherche en santé du Canada

NSERC CRSNG

SSHRC CRSH

Curation Outreach Preservation Discovery

Training Policy Development

Sensitive Data

https://static1.squarespace.com/static/576bfb4b2e69cfe8cbce5a20/5786be9eff7c509730365440/578a66eabebafb8c81408a37/1468688108359/172+-+Canoe+Portage+Sign+T-Shirt+-+Canada+Clothing.jpg?format=500w
https://cdn0.iconfinder.com/data/icons/wthr-color/32/693734-tide-512.png

Next Steps

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

- Build a Community of Data Curators (CEG)
- Add Geospatial discovery features for FRDR & Dataverse
- Launch “RDM 101” training modules (TEG)
- Improve tools for “active” RDM (CCF/Radiam)
- Offer FRDR & Dataverse as national repository options
- Transition to ‘new organization’ under ISED funding

More to come!

New Business Models

CIHR Training Modules

SSHRC DMP Workshop

COPPUL MOU

CSA -Life Sciences

Website Launch

Compute Canada/FRDR

DMP Assistant

First Portage Director

CARL Approval

2016

15

2015

2014: Project ARC

DVN Report

BCI MOU

Sensitive Data WG

Regional Library Consortia

Portage Day - Toronto

Institutional Strategy WG

Second Portage Director

FRDR – Limited Production

Portage Day - Montreal

Formative Assessment

LCDRI

CAUL MOU

Dataverse North

OCUL MOU

New Governance Model

115

2018

98

2017

42

More to come!

Colleges

Non-CARL Libraries

International Outreach

Portage/Tri-Agency Tour

Preservation White Paper

CARL Directors Support

2019

2018

Governance

Platforms/Services

Outreach

Agreements

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

Questions?

portage

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

portagenetwork.ca

Jeff Moon, Portage Director
portagedir@carl-abrc.ca