

portage

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



Portage Network: Update on Progress and Transition

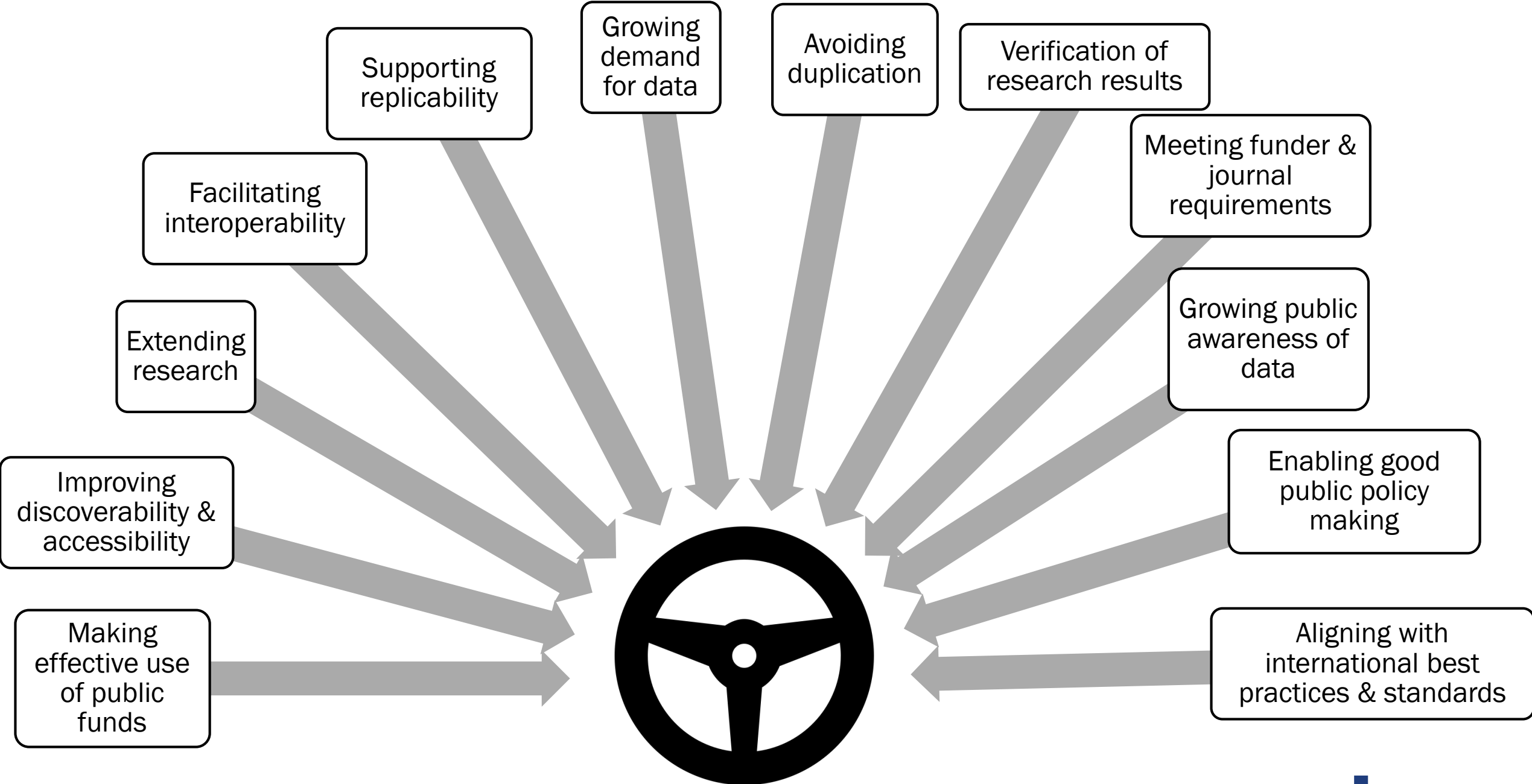
Jeff Moon, Director, Portage
NDSF Summit, Ottawa | 5 February 2020

Agenda

- RDM Drivers
- FAIR Principles & Portage
- RDM in Transition



RDM Drivers



RDM Drivers

What keeps me up at night . . .



Data

Commercial interests





Preventing data loss



. . . it can happen to you!

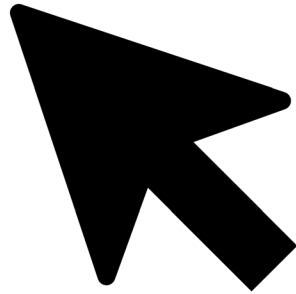


FAIR Principles & Portage

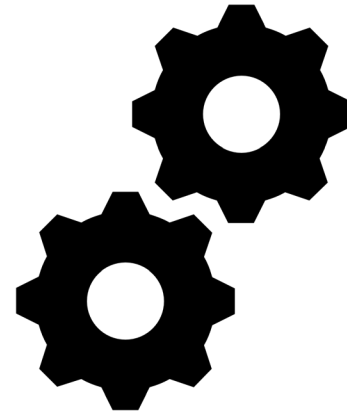
Fair Principles



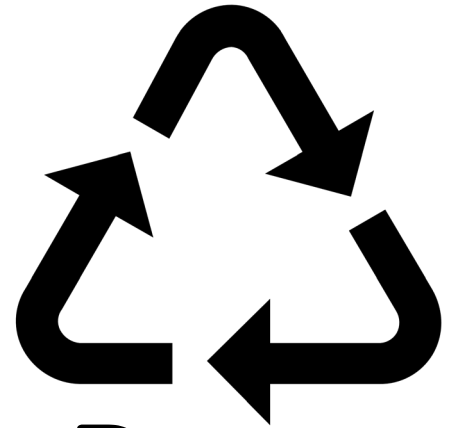
Findable



Accessible

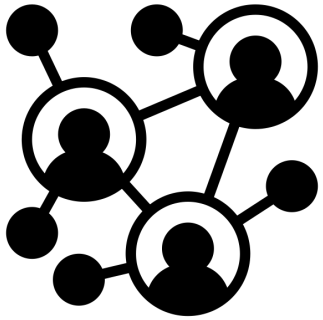


Interoperable

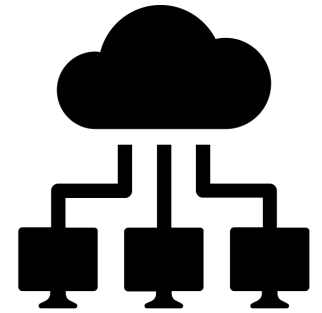


Reusable

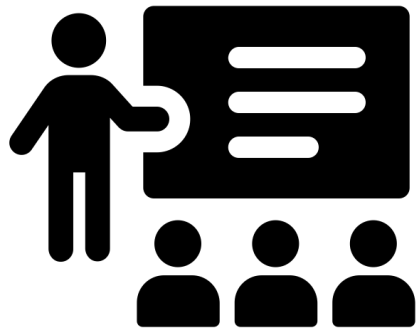
A set of principles to ensure that data are shared in a way that enables and enhances reuse by humans and machines



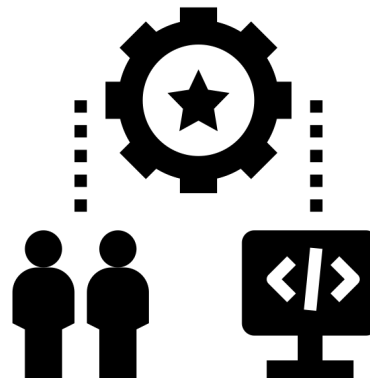
Network of Experts



Infrastructure Platforms



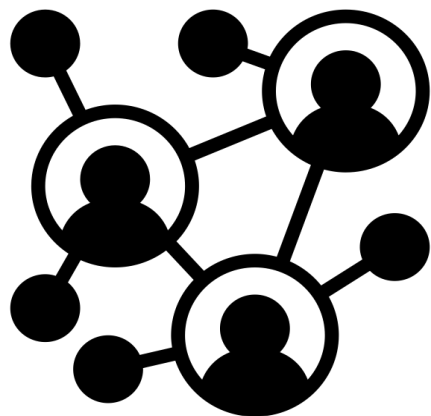
Training



Services



Tools



Network of Experts

Expert Groups

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



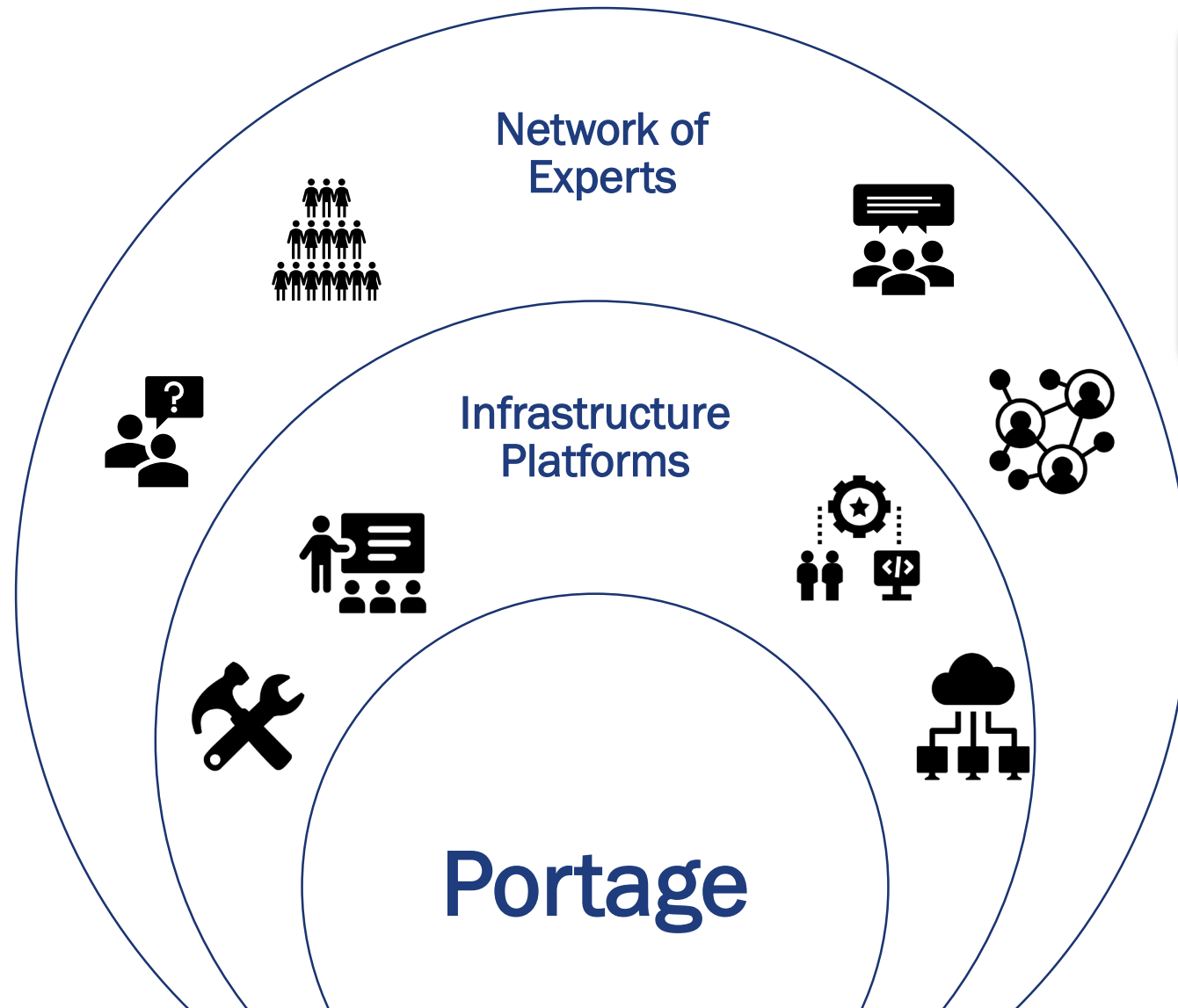
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

130+ Experts

60+
Organizations

Working Groups

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



Make RDM best
practices
possible, then
make them
easy...

Fair Enablers



Findable



Accessible

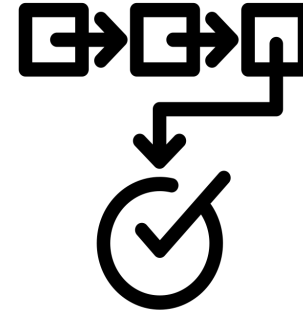


Interoperable

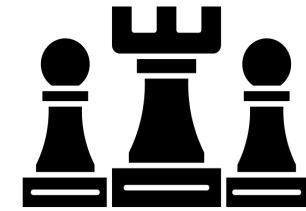


Reusable

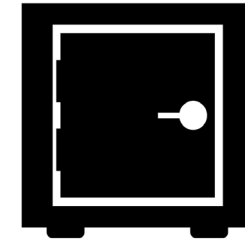
Funder
Policies



Data Management
Plans (DMPs)



Institutional
Strategy

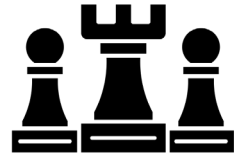


Deposit



SSHRC  CRSH

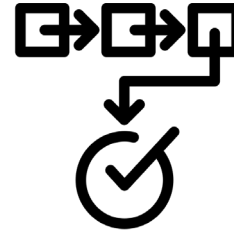
Portage actions related to Tri-Agencies' policy...



Institutional
Strategy



RDM
Strategy
Template



Data Management
Plans (DMPs)



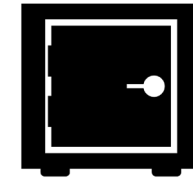
DMP
Assistant

National, online, bilingual, Data
Management Planning Tool

DMP Assistant

Sign In

Sign Up



Deposit



National,
Multi-disciplinary
Repository
Options

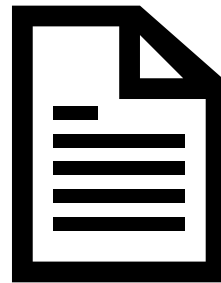
portage

Portage RDM Strategy Template

Strategy components

1. Raise Awareness
2. Assess institutional readiness
3. Formalize RDM practices
4. Define a roadmap

+ associated guidance



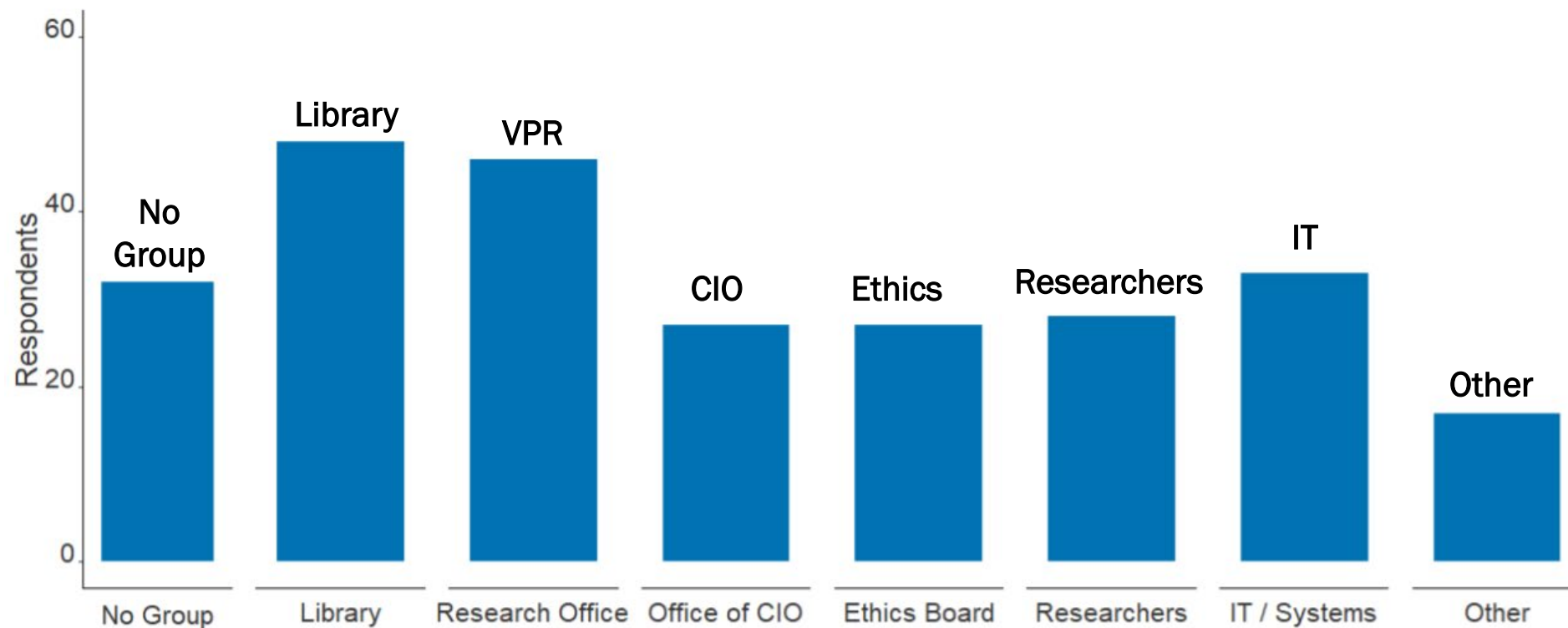
RDM Strategy Template

Institutional RDM Strategy Survey

- 6 questions
- 88 responses
- 63 institutions
- June 2019

Institutional RDM Strategy Survey:

Representation on campus working groups



Institutional RDM Strategy Survey:

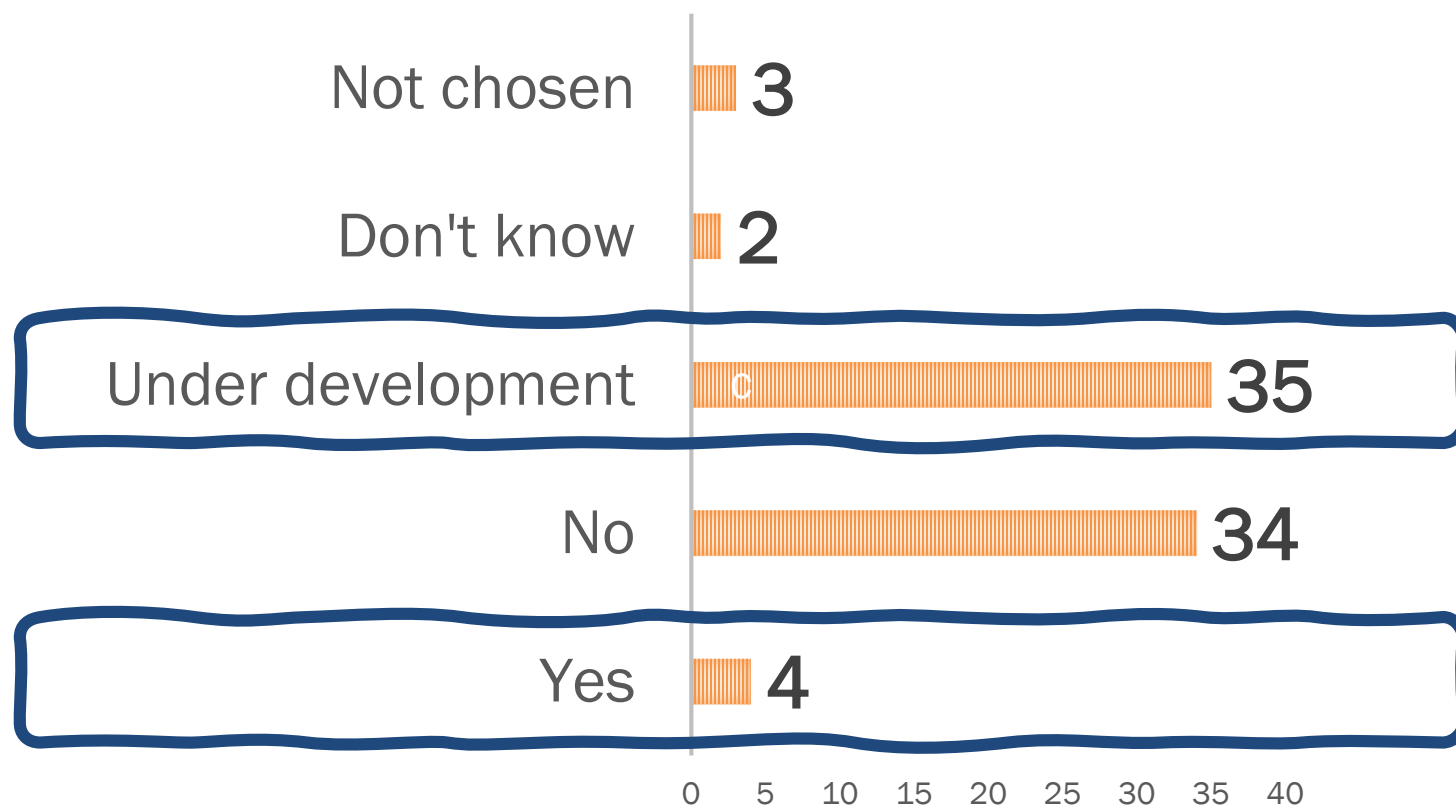
Resources being used to develop RDM Strategy



Institutional RDM Capacity Survey

- 27 questions
- 78 institutional responses
- September-October 2019

Institutional RDM Services Capacity Survey



Does your institution currently have a research data management strategy?

Figure 2. Level of RDM strategy development by institution. (n=78). Q14a “Does your institution currently have a research data management strategy?”



DMP Assistant

- National
- Online
- Bilingual
- Data Management Planning Tool

Transition to 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 or ‘active’ DMPs
- Supports DMPs as “living documents”
- Supports use of ORCID IDs
- Includes APIs for creating and sharing DMPs
- Supports researchers ‘opting in’ to share their DMP

Transition to 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 or 'active' DMPs
- Supports DMPs as “living documents”
- Supports use of ORCID IDs
- Includes APIs for creating and sharing DMPs
- Supports researchers 'opting in' to share their DMP

DMP-related initiatives

Portage is developing a number of discipline-specific **DMP Exemplars** and a **DMP-specific repository** to enhance discovery and sharing.

Customized **DMP Templates**

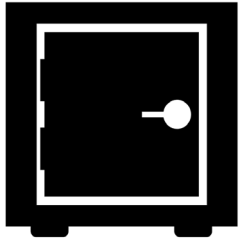
By institution

By discipline

By use case

→ Indigenous data template





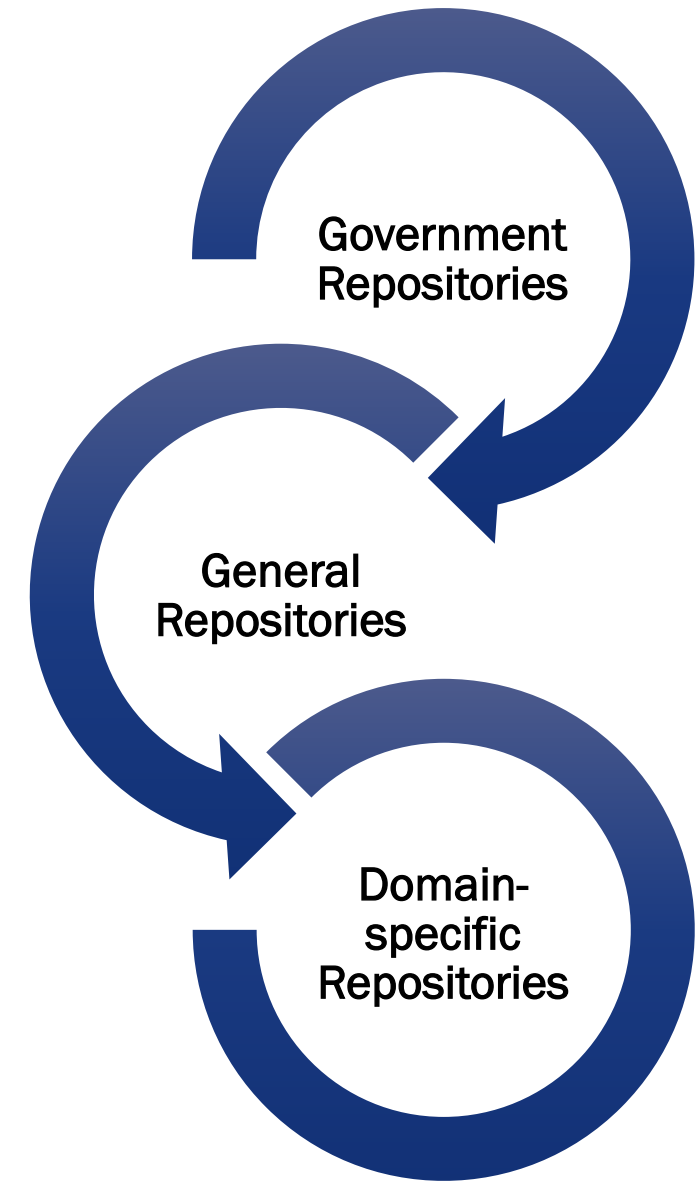
National,
Multi-disciplinary
Repository Options



Federated Research Data Repository

Dataverse
North

National Instance of Dataverse



portage

Institutional RDM Services Capacity Survey

What research data management infrastructure does your institution have access to?

Researchers use disciplinary/multidisciplinary data repositories

37

Researchers use regional or consortial data repositories

23

Researchers use journal repositories

27

Researchers use institutional data repository

27

Other

23

Don't know

17

0 20 40 60 80
Number of Institutions

Almost half reported use of disciplinary/multidisciplinary repositories

Figure 6b. Data repository access. Institutions (n=78) could select all that applied. Q10 "What research data management infrastructure does your institution have access to? Check all that apply."

Dataverse

The logo icon for Dataverse, consisting of three orange circles of increasing size connected by lines, forming a stylized 'D' or a network structure.

portage

Dataverse



Scholars
Portal
Dataverse

Institutions

51

Dataverses

450+

Datasets

2,307

Files

29,862

Downloads

178,642

Nov 12, 2019



University of British Columbia licensed
Simon Fraser University
University of Northern British Columbia



UAL
Dataverse



DALHOUSIE
UNIVERSITY
Dataverse



DATVERSE
RESEARCH DATA REPOSITORY

portage



Regional Partnerships

- Agreement with **BCI** to serve 19 university libraries in Quebec
- Working on agreements with **COPPUL** and **CAUL**

Policy Development

Engaging with Portage's
Dataverse North Working Group to develop policy

Funding

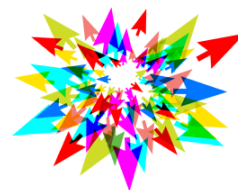
Transitional funding from Portage will support **storage expansion** to subsidize costs for institutions



Federated Research Data Repository

Federated Research Data Repository

A scalable, federated platform for digital research data management and the discovery of Canadian research data



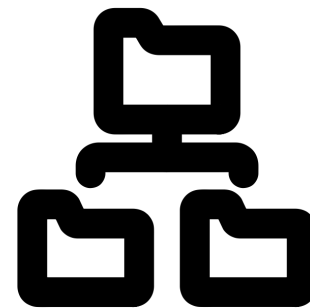
compute
canada

calcul
canada

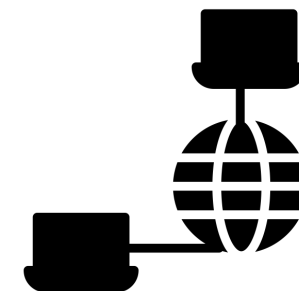
portage



Big data capable
& scalable



Retains file
hierarchies



Geographically
distributed

portage

Federated Research Data Repository

A scalable, federated platform for digital research data management and the discovery of Canadian research data



>40 partnering
research
groups

>100 datasets
(9+ TB)

Production launch in Fall
2020

Informational and training
sessions will be provided
ahead of launch



Federated Research Data Repository



FRDR Transitioning to Full Production

2
National
Hosting
Sites

Data deposit for
all researchers
(PIs) and
librarians, or their
designates, from
Canadian
postsecondary
institutions

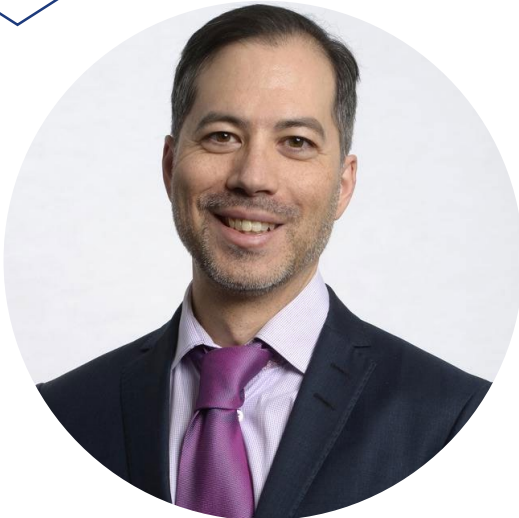
Access to **curated
repository storage**
without cost up to
a certain
threshold;
more storage
possible if needed

National Discovery Layer

“There needs to be a
central clearinghouse of
publicly available datasets
in Canada”



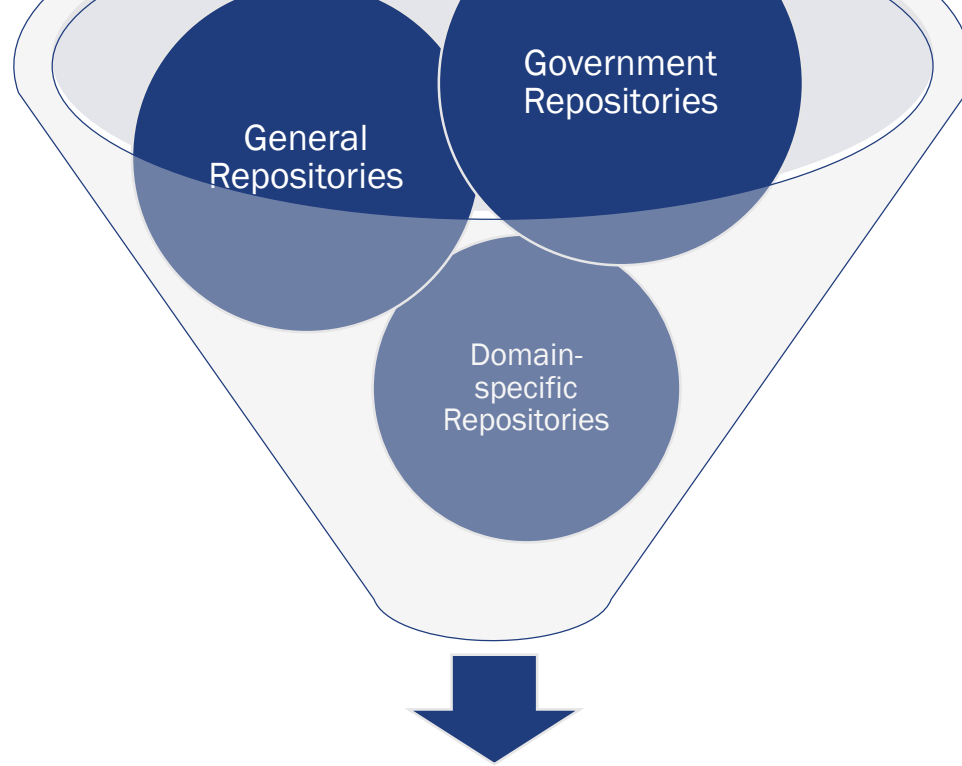
Globe & Mail
28 January, 2019



Fuyuki Kurasawa

Director, Global Digital Citizenship Lab
York University





Metadata harvested to FRDR

National Discovery Layer



Improve discovery
of Canadian
research (meta)data

Break down
repository siloes

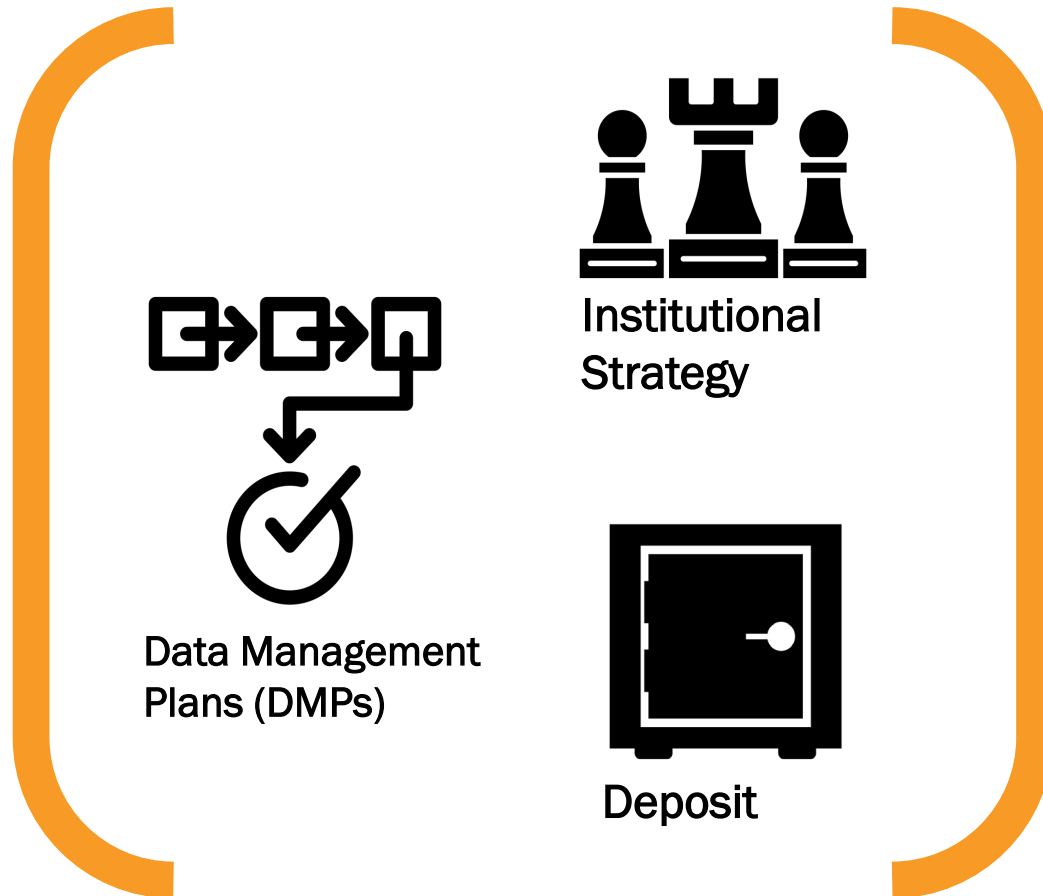
Drive traffic to
existing repository
sites

Create
interoperability
between Canadian
and international
platforms



portage

Core activities of Portage to date...



National Discovery Layer



portage

Fair Enablers



Findable



Accessible



Interoperable



Reusable

Persistent Identifiers (PIDs)

- DOIs for data or journal publications



CRKN & CARL Portage: The DataCite Canada Consortium

DataCite Blog

Good news for the new year: the DataCite Canada Consortium will launch on January 1, 2020

December 18, 2019 by Helena Cousijn • 3 min read

<https://doi.org/10.5438/agr4-ee34>

Following the [updates to the DataCite member model](#) earlier this year, we are working with several organizations on establishing DataCite consortia. We are therefore very excited to announce that the DataCite Canada Consortium will [launch](#) on January 1, 2020!

For the past ten years, DOIs in Canada were provided through the National Library of Medicine. When we started conversations about forming a DataCite consortium, the Canadian Research Knowledge Network (CRKN) might be willing to provide a collective negotiating body for academic libraries in Canada. As part of this effort, CRKN leads the Canadian ORCID consortium, and provides tools, and services, CRKN leads the Canadian ORCID consortium, and provides a community around persistent identifier (PID) infrastructure.

[Visit: https://stats.datacite.org/](https://stats.datacite.org/)

58
Organizations

> 425,000
DOIs

Fair Enablers



Findable



Accessible



Interoperable



Reusable

Persistent Identifiers (PIDs)

- DOIs for data or journal publications
- ORCID for individuals



ORCID-CA Consortium

The ORCID Consortium in Canada



17 member multi-stakeholder working group

Fair Enablers



Findable



Accessible



Interoperable



Reusable

Persistent Identifiers (PIDs)

- DOIs for data or journal publications
- ORCID for individuals
- Research Organization Registry (ROR) for organizations



Fair Enablers



Findable



Accessible



Interoperable



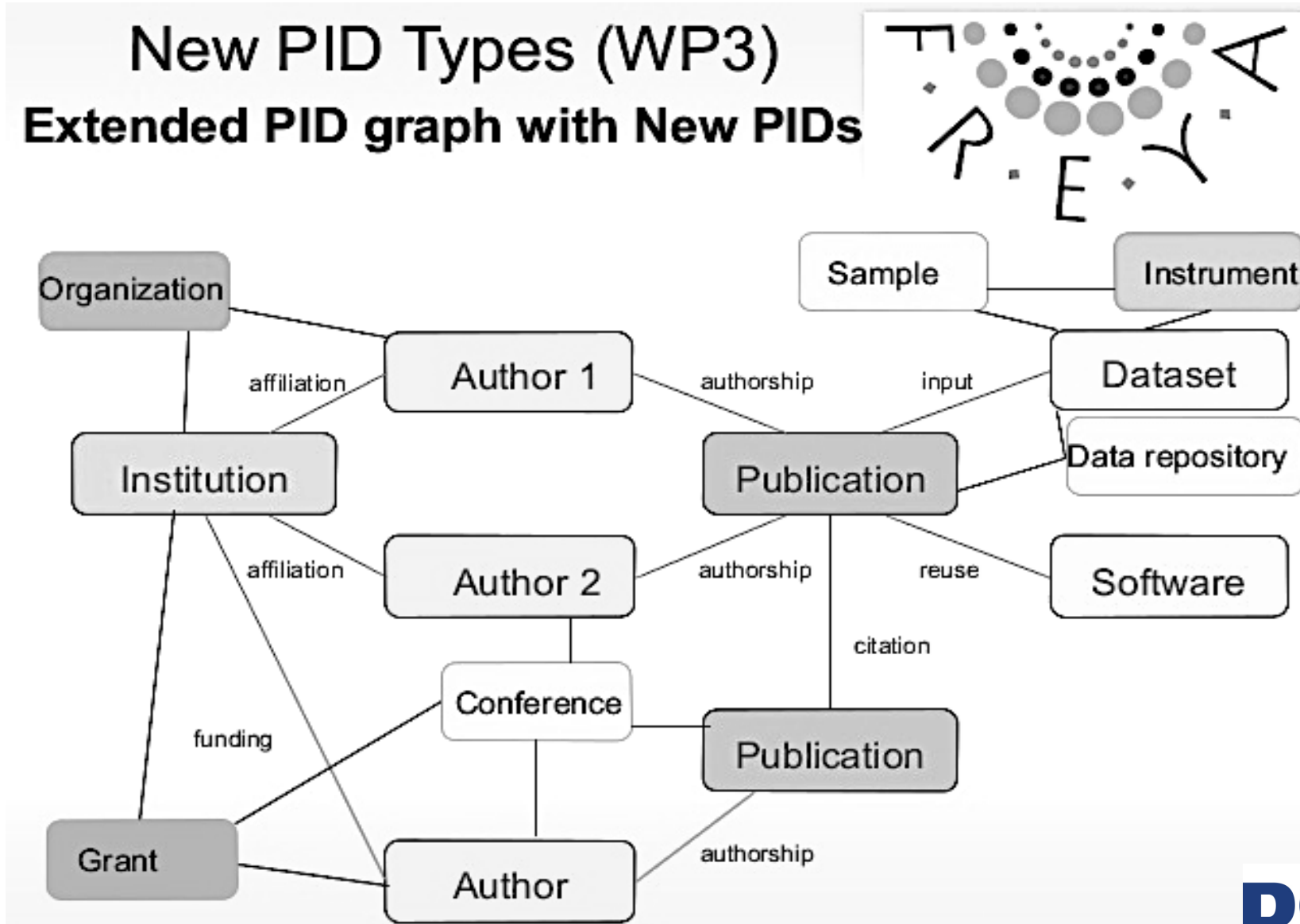
Reusable

Persistent Identifiers (PIDs)

- DOIs for data or journal publications
- ORCID for individuals
- Research Organization Registry (ROR) for organizations
- RAiD (Research Activity Identifiers) for projects



Other Potential PIDs



Fair Enablers



Findable



Accessible



Interoperable



Reusable

Repository Certification



RDC  **DRC**
Research Data Canada – Données de Recherche Canada

Best Practices

COAR
Confederation
of Open Access
Repositories

portage

Fair Enablers



Findable



Accessible



Interoperable



Reusable



OCLC
FAST

Subject Headings

(Faceted
Application of
Subject
Terminology)

Metadata, Controlled Vocabularies
& Discovery

Improving &
expanding
metadata
harvesting
from Canadian
Repositories

Improving
Geospatial
discovery
through
GEODISY
Project

Fair Enablers



Findable



Accessible



Interoperable



Reusable

Interoperability



FAIRsFAIR

Fostering Fair Data Practices in Europe

FRDR selected to receive support from FAIRsFAIR to improve international interoperability

portage

Fair Enablers



Findable



Accessible



Interoperable



Reusable

Sensitive Data

- ✓ Glossary of 'sensitive data' terms
- ✓ Defining risks related to roles of individuals
- ✓ Deposit-friendly text for ethics, informed consent
- ✓ Data access agreements
- ✓ **Research Data Risk Matrix**



Fair Enablers



Findable



Accessible



Interoperable



Reusable



SIMON FRASER
UNIVERSITY

Sensitive Data

- ✓ **FRDR-aligned project** to support sensitive data
- ✓ Exploring viability of zero-knowledge encryption in an RDM context
- ✓ Developing tools to encrypt packages at time of deposit
- ✓ Supporting metadata-only review and discovery of sensitive datasets.



Fair Enablers



Findable



Accessible



Interoperable



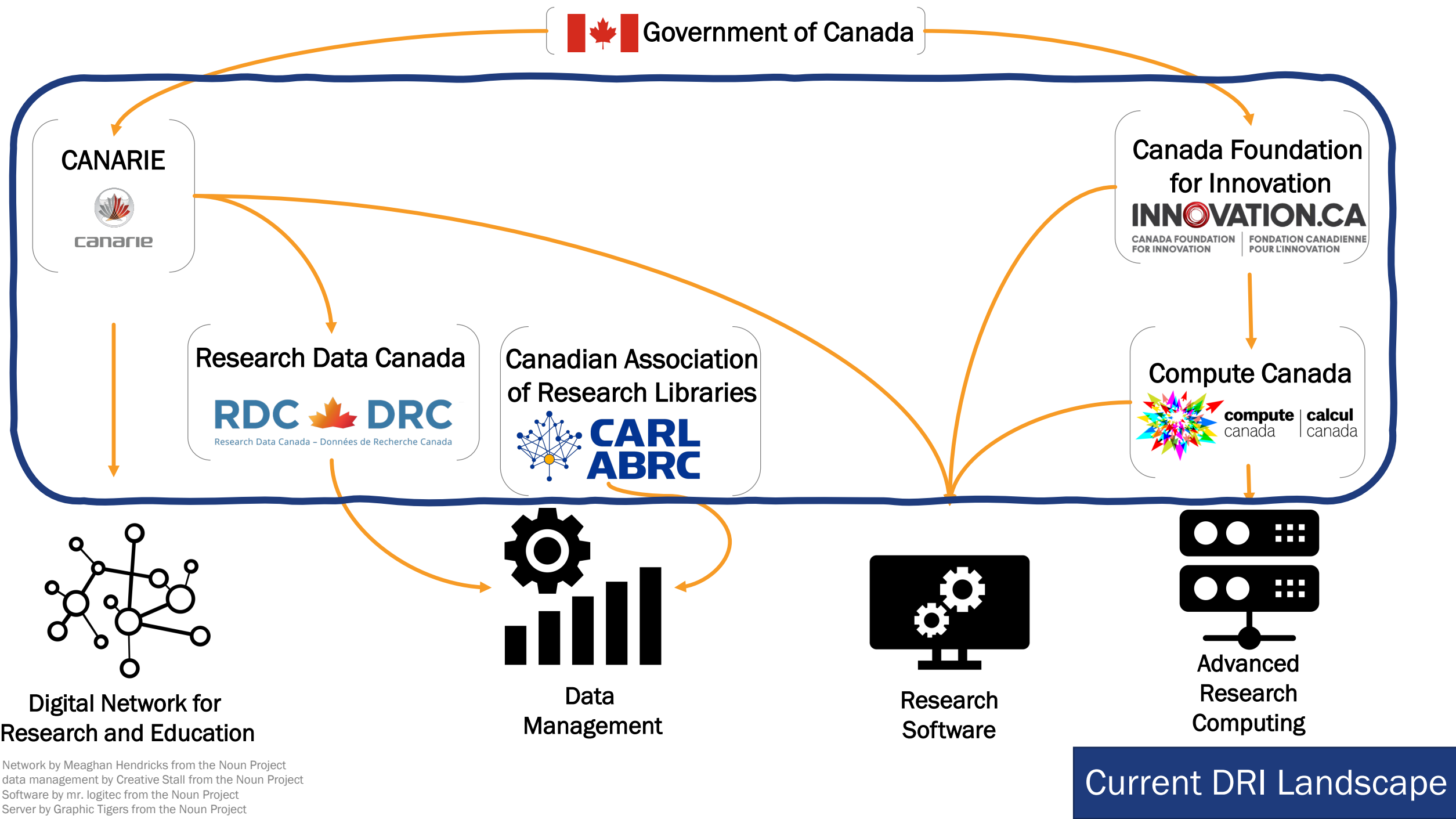
Reusable

Other areas of activity

- Preservation workflow
- National network of curation support
- Training and outreach activities
- Governance – Dataverse & FRDR
- Increased engagement with domains



RDM in Transition



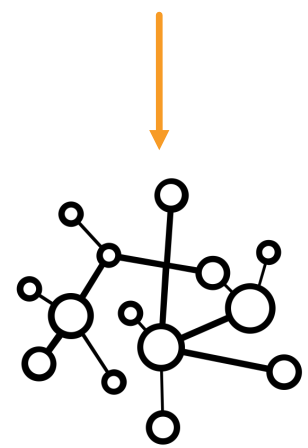
Network by Meaghan Hendricks from the Noun Project
data management by Creative Stall from the Noun Project
Software by mr. logitec from the Noun Project
Server by Graphic Tigers from the Noun Project

Current DRI Landscape

 Government of Canada

CANARIE

canarie

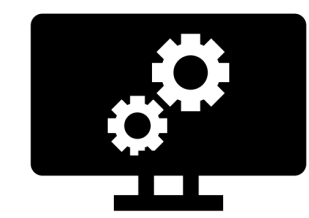


Digital Network for
Research and Education

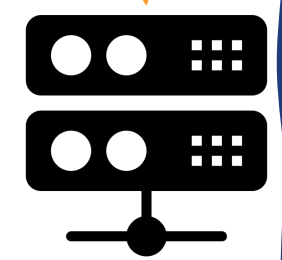
New DRI Organization



Data
Management



Research
Software



Advanced
Research
Computing

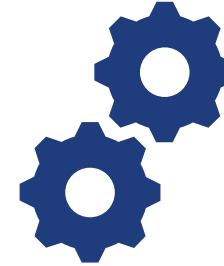
Moving toward a new DRI organization

Key transition priorities for Portage



National Data Stewardship Support

Providing coordinated support at the
national level



RDM Platform Support and Development

Develop and operationalize national,
collaborative RDM platforms and
services

National Data Stewardship Support



Erin Clary
Preservation
Coordinator



Beth Knazook
Preservation
Coordinator



Kelly Stathis
Discovery
&
Metadata
Coordinator



Melanie Parlette-Stewart
Training
Coordinator



Caitilin O'Hare
Communications
Officer



Veronica Amirault
Project
Manager



Jennifer Abel
Project Officer



Lee Wilson
Service Manager



Nora



Jeff Moon
Director

RDM Platform Support & Development



Partners



DMP Assistant

[Sign In](#)

[Sign Up](#)



FRDR Archival Storage



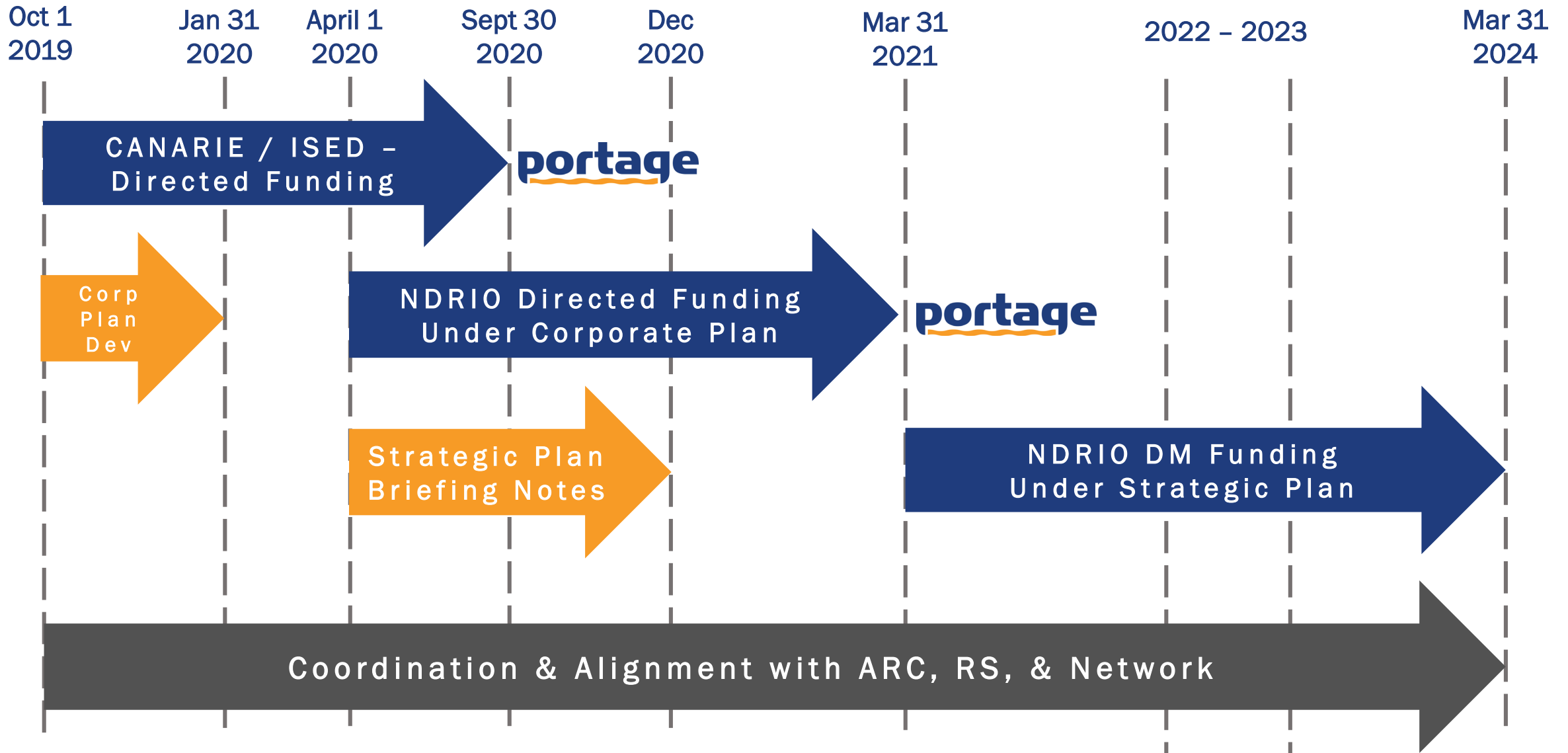
FRDR Project Lead



FRDR Host Sites



DM Timeline for transition to NDRIO



Looking Forward

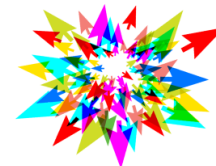
- Improving discovery
- Developing & deploying a national RDM training strategy
- Growing a national network of data curation support
- Supporting the evolution of national preservation services
- Addressing issues related to sensitive data
- Working closely with domains to address discipline-specific issues
- Advancing repository certification efforts in Canada

Portage

“I was at the tracks recently and won a tidy sum on a horse named Portage”



Simon Hodson
Executive Director
CODATA



compute | **calcul**
canada | canada

portage

A photograph of a horse race in progress. Several jockeys on horses are running on a grass track. In the background, there is a large, green hill under a cloudy sky. The text "CARL Library Directors (and others) decided to back Portage..." is overlaid on the image in white, sans-serif font.

CARL Library Directors
(and others) decided
to back Portage...

...and this backing has paid off well in serving
the RDM needs of researchers and institutions

Thank you to our partners

CARL Portage is supported by directed funding from Innovation, Science & Economic Development Canada (ISED), flowing through CANARIE



Innovation, Science and
Economic Development Canada
Innovation, Sciences et
Développement économique Canada



canarie



compute
canada | calcul
canada

portage



Network of Experts



UNIVERSITY OF
TORONTO

SciNet

UNIVERSITY OF
WATERLOO



UNIVERSITY OF
SASKATCHEWAN



SIMON FRASER
UNIVERSITY

OCUL

Ontario Council of
University Libraries



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

Questions?