Portage Update

OLA RDM Day
Toronto
January 30, 2018

Jeff Moon, Director, Portage
What is the Portage Network?

“Portage is a national, library-based research data management network that coalesces initiatives in research data management to build capacity and to coordinate activities better.”
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**
Drivers in the RDM Ecosystem

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 (RDM)
Portage Network of Experts: Expert Groups

Data Management Planning
- DMP Assistant 2.0
- Domain-specific DMPs

Curation
- Briefing Notes
- Data Curation White Paper

Data Discovery
- Core metadata standards
- New FRDR Discovery Service WG

Preservation
- Briefing Notes
- Preservation White Paper

Training
- Resources on portagenetwork.ca
- Training sessions/workshops

Research Intelligence
THE PORTAGE NETWORK is dedicated to the shared stewardship of research data in Canada through:

- Developing a national research data culture
- Fostering a community of practice for research data
- Building national research data services and infrastructure

Launched in 2015 by the Canadian Association of Research Libraries, Portage works within the library community to coordinate expertise, services, and technology in research data management, seeking to collaborate with other research data management stakeholders.

Research data culture represents widely shared values and principles for digital data management.

A community of practice for research data consists of stakeholders working collaboratively to ensure data are accessible to address complex research issues.

News

- Registration Now Open! Portage and RDM in Canada - Toronto, Jan. 30 2018 - December 21, 2017
- Call for Participation: Portage Data Management Planning Expert Group - November 21, 2017
- Save the Date! Portage and RDM in Canada - Toronto, Jan. 30, 2018 - November 16, 2017
"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.

1. Save your raw data in original format
   1.1. Don’t overwrite your original data with a cleaned version.
   1.2. Protect your original data by locking them or making them read-only.
   1.3. Refer to this original data if things go wrong (as they often do).

2. Backup your data
   2.1. Use the 3-2-1 rule: Save three copies of your data on two different devices, and one copy off-site.
   2.2. Do not backup or store sensitive data on a commercial cloud (Dropbox, Google Drive, etc.).

3. Describe your data
   3.2. Human Friendly: Describe your variables, so your colleagues will understand what you meant. Data without good metadata is useless. Give your variables clear names.
   3.3. Use "NA" for missing data (computer programs don’t like blank cells).
   3.4. Convert your data to open, non-proprietary formats.
   3.5. Name your files well with basic metadata in file names.

4. Process your data
   4.1. Make each column a variable.
   4.2. Make each row a case.

5. Archive and preserve your data
   5.1. Submit final data files to a repository alongside a persistent identifier.

---

External Training Resource Library
<table>
<thead>
<tr>
<th>Portage Network of Experts: Expert Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Management Planning</strong></td>
</tr>
<tr>
<td>DMP Assistant 2.0</td>
</tr>
<tr>
<td>Domain-specific DMPs</td>
</tr>
<tr>
<td><strong>Curation</strong></td>
</tr>
<tr>
<td>Briefing Notes</td>
</tr>
<tr>
<td>Data Curation White Paper</td>
</tr>
<tr>
<td><strong>Data Discovery</strong></td>
</tr>
<tr>
<td>Core metadata standards</td>
</tr>
<tr>
<td>New FRDR Discovery Service WG</td>
</tr>
<tr>
<td><strong>Preservation</strong></td>
</tr>
<tr>
<td>Briefing Notes</td>
</tr>
<tr>
<td>Preservation White Paper</td>
</tr>
<tr>
<td><strong>Training</strong></td>
</tr>
<tr>
<td>Resources on portagenetwork.ca</td>
</tr>
<tr>
<td>Training sessions/workshops</td>
</tr>
<tr>
<td><strong>Research Intelligence</strong></td>
</tr>
<tr>
<td>Environmental scan/Roadmap</td>
</tr>
<tr>
<td>Identifying gaps and priorities</td>
</tr>
</tbody>
</table>
Portage Network of Experts: Working Groups

6 Expert Groups
- Data Management Planning
- Curation
- Data Discovery
- Preservation
- Training
- Research Intelligence

Working Groups
- Dataverse North WG
- FRDR Service Model WG
- Ethical Treatment of Sensitive Data WG
- FRDR Discovery Service WG
- Institutional RDM Strategies WG

Institutional Strategy

DMPs

Deposit

Metadata standards
Ethics, Privacy, IP
RDM Best Practices
Service Models
and more…
Portage Network of Experts

Commitments of Portage Volunteers

- One Group: 90+
- Two or more: 16

Portage Volunteers by Region

- Atlantic: 39
- Quebec: 28
- Ontario: 12
- West: 8

Portage Volunteers by Organization

- CARL Universities: 24
- Non-CARL Universities: 5
- Organizations: 6
A True Network of Experts!

Credit: Lee Wilson
Building a Community of Practice

Institutions
- Liaison Librarians
- IT Specialists
- University Research Offices
- Ethics Boards
- Other stakeholders

Network of Experts
- RDM Specialists/Librarians
Building a Community of Practice

Network of Experts
- RDM Specialists/Librarians

Institutions
- Liaison Librarians
- IT Specialists
- University Research Offices
- Ethics Boards
- Other stakeholders

Researchers
- Academic
- Government
- Others

Network of Expertise Layer

Institutional Layer

Researcher Layer
Building a Community of Practice

Network of

Researchers

Institutions

Network of Experts

Infrastructure Platforms

Liaison Librarians
IT Specialists
University Research Offices
Ethics Boards
Other stakeholders

RDM Specialists/Librarians

DMP Assistant
Federated Research Data Repository
Dataverse

Academic
Government
Others

Educators

Infrastructure Platform

Institutional Layer

Network of Expertise Layer

Investigator Layer
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
Local Stakeholders

Libraries

Researchers

Graduate Studies

IT

Ethics

Research Services
Regional Stakeholders

Canada’s four regional library consortia

COPPUL
Council of Prairie & Pacific University Libraries

OCUL
Ontario Council of University Libraries

BCI
Bureau de Coopération Interuniversitaire

CAUL/CBUA
Digital Research Ecosystem

Adapted from CFI DRI submission
Digital Research Ecosystem

Adapted from CFI DRI submission
Digital Research Ecosystem

Adapted from CFI DRI submission
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
Data Management Planning

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

Research Data Repositories

Federated Research Data Repository

https://portagenetwork.ca/frdr-dfdr

Portage partnership with Compute Canada

Regional & Institutional Instances of Dataverse

Portage Dataverse North Working Group

Institutional Strategy  DMPs  Deposit

Institutional Strategy  DMPs  Deposit
Research Data Repositories & Portage

Federated Research Data Repository

Main Features:
- ‘Big Data’ upload/download
- National Discovery Platform
- National Repository Option
- Provides Preservation Pipeline

https://portagenetwork.ca/frdr-dfdr

Dataverse

Main Features:
- Mature platform
- Widely adopted
  - Scholars Portal
  - Abacus
  - Dal/CAUL
  - Others…
- User-friendly
- Brandable
- Possible National Repository Option

Portage Dataverse North Working Group
Federated Architecture

Council of Prairie & Pacific University Libraries

Ontario Council of University Libraries

BCI
BUREAU DE COOPÉRATION INTERUNIVERSITAIRE

CAUL/CBUA
Federated Architecture

Preservation processing via Archivematica

National Discovery Layer

Metadata harvested into FRDR

FRDR

Regional/Institutional instances of Dataverse

Regional/Institutional Repositories

Discipline-specific Repositories

Research data & metadata
Deposit & Download

‘Big Data’ (& small) deposit & download
(a national repository option)
Federated Architecture - Storage

Active Storage: Research-level storage that facilitates actual research process.

Archival Storage: Network of Preservation Storage Providers.

Repository Storage: Dataverse, FRDR, & Other repositories.

Preservation processing via Archivematica.

National Discovery Layer.

Network of Preservation Storage Providers.

Regional / Institutional instances of Dataverse.

Regional / Institutional Repositories.

Discipline-specific Repositories.

Metadata harvested into FRDR.

Research data & metadata Deposit & Download.

FRDR:

Portage.

‘Big Data’ (& small) deposit & download using FRDR.

Active Storage.

Research-level storage that facilitates actual research process.
Data Storage Continuum

Research Lifecycle from a storage perspective

**ACTIVE STORAGE**
- Controlled
- Working copy Short-term for duration of project
- Completing research

**REPOSITORY STORAGE**
- Open (as appropriate)
- Dissemination copy Medium-term, beyond duration of project
- Discovery & Access

**ARCHIVAL STORAGE**
- Open (as appropriate)
- Preservation copy Long-term
- Disaster recovery/Copy of last resort

**ACCESS:**

**STORAGE PURPOSE:**

**USE:**
Data Storage Continuum

ACTIVE STORAGE
- Institutional network drives
- Laboratory computers
- Laboratory equipment
- High performance computing clusters

REPOSITORY STORAGE
- Discipline-specific Repositories
- Regional & Institutional Repositories
- Regional & Institutional instances of Dataverse

ARCHIVAL STORAGE
- Distributed Network of Preservation Service Providers PSPs
- Archival Information Packages AIPs
- Preservation (archival) processing via ARCHIVEMATICA

Metadata harvested into FRDR

Portage

Federated Research Data Repository (FRDR)

National Discovery Layer
And there are other RDM activities that happen along with ‘storage’
Portage Outreach: with Tri-Agencies

Jan 2018:
Vancouver & Calgary

Sept 2017:
Montreal

Feb 2018:
Halifax

Jan 2018:
Toronto
Portage Outreach: Fall 2017

**September:** Portage/RDA RDM Day

**November:** ACCOLEDS – RDM at DLI Training, Thompson Rivers U, Kamloops

**December:** RDM at Ontario Regional DLI Training, Ryerson U, Toronto

**December:** Portage update to CODATA Canada [webinar]

**December:** Coalition for Networked Information (CNI) Conference (Donna Bourne-Tyson, Lee Wilson, Corey Davis), Washington, DC
February: 13th Annual International Digital Curation Conference (IDCC), Barcelona, Spain

February: Meetings with LAC and NRC regarding RDM

March: CODATA Conference, Gottingen, Germany

March: Research Data Alliance (RDA) Conference, Berlin, Germany

April: BCNET 2018, Vancouver, BC

May: IASSIST 2018, Montreal, QC

June: Atlantic Provinces Library Association, Fredericton, NB
So Portage has been busy!
Highlights 2015 to Fall 2017
So Portage has been busy!
Highlights Fall 2017 to Date

Continued Excellent work by Expert & Working Groups

Several New Targeted Working Groups

FRDR in Limited Production

MOUs with Regional Library Consortia, and regular communication

Outreach to stakeholders in collaboration with Tri-Agency

International outreach at CNI, IDCC, CODATA, RDA, & IASSIST

Contributions to LCDRI and ISED reports and discussions
Questions?

portagenetwork.ca

Jeff Moon, Portage Director
moonj@queensu.ca