The Portage Network
the shared stewardship of research data

Chuck Humphrey, Director of Portage
January 26, 2017
How statistics lost their power – and why we should fear what comes next

William Davies
The Guardian
19 January 2017
Alternative Facts

Kellyanne Conway
NBC’s “Meet the Press”
22 January 2017
Facts are derived from data
Research data management is important!
The Portage Mission

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

● fostering a national research data culture
● developing a community of practice for research data, and
● building national research data services and infrastructure.
Research data culture

● Values and norms that provide an understanding of research data in our society.
  ○ Data culture of use → evidence-based actions
  ○ Data culture of sharing → allowing others access to your research data
  ○ Data culture of stewardship → taking responsibility for the long-term access to your research data
Tri-Agency Statement of Principles

- The Tri-Agency Statement of Principles on Digital Data Management is an example of values dealing with research data use, sharing, and stewardship.

The agencies believe that research data collected with the use of public funds belong, to the fullest extent possible, in the public domain and available for reuse by others. They also strongly support the creation of a robust and efficient environment for data stewardship in Canada and internationally. [p. 2]
Tri-Agency Statement of Principles

- The statement further describes practices that support these values:
  - Data management planning
  - Working within legal and ethical obligations
  - Adherence to standards
  - Secure digital practices
  - Providing quality metadata
  - Preserving, retaining, and sharing
  - Timely sharing
  - Citation and attribution
  - Efficient and cost effective practices
Data stewardship

Who is responsible during and after a research project?
Institutional shared-stewardship

Libraries

Researchers

Graduate Studies

Individuals, Groups and Services

IT

Ethics

Research Services
Community of practice

- Portage is working with other research data management stakeholders in developing a broader community of practice dealing with the use, sharing, and stewardship of research data.
- Portage is working specifically with library colleagues addressing the culture of data stewardship.
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.

Research data management activities span sectors, domains, and jurisdictions,
Collaborators in Research Data Management

Directions for Research Data Management (RDM)
- RDM multi-sector stakeholders’ forum
- RDM policy & best practices directions
- International RDM voice for Canada
- Strategic & competitive RDM investments

Facilitating the development of a best practices, sustainable national RDM ecosystem.

Operations for Research Data Management (RDM)
- RDM community of practice
- RDM services & infrastructure platforms
- National RDM network of expertise
- Coordination of library-based RDM initiatives

Developing and delivering RDM services and resources to Canadian universities and partners.

Supported by:

[Logos for Canarie and CARL ABRC]
Portage data services & infrastructure

- I would now like to focus on Portage’s contributions to research data services and infrastructure in Canada.
The Portage service model

A federated vision:

• A federated model has been adopted by CARL that leverages coordinated RDM contributions from libraries and partnering stakeholders.

The Portage service model

A federated vision:

- This federated structure consists of a Network of Expertise and platforms to support data management planning, data curation, data preservation, and data discovery.
- Experts are located in individual institutions but are coordinated to provide services and advice from a network level.
Network of Expertise

- The six expert groups
  - Data Management Plans (DMPEG)
  - Curation (CEG)
  - Preservation (PEG)
  - Discovery (DEG)
  - Training (TEG)
  - Research Intelligence (RIEG)

- Two working groups are functional
  - Collection development policies for data repositories
  - Metadata for discovery
Network of Expertise

- Members of expert groups agree to two year commitments; working group members serve for three to six months on task-specific activities.
- The Chairs of the Expert Groups are members of the Council of Chairs, which is responsible for coordinating overlap among the groups, for planning short and longer term goals of Portage, and for identifying budgetary items associated with each group.
## Network of Expertise

<table>
<thead>
<tr>
<th>Expert Group</th>
<th>Number of Individuals</th>
<th>Number of Unique Institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMPEG</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>CEG</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>PEG</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>DEG</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>TEG</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>RIEG</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>44†</strong></td>
<td><strong>19</strong></td>
</tr>
</tbody>
</table>

† Seven of the members serve on multiple expert groups, including the Portage Director, who is an ex officio member of all groups. Thirty-three individuals are currently serving in the 44 positions.
## Discovery working groups

<table>
<thead>
<tr>
<th>Regional consortium</th>
<th>Metadata WG</th>
<th>Collection Development WG</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAUL (Atlantic)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>BCI (Quebec)</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>OCUL (Ontario)</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>COPPUL (West)</td>
<td>3</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>11</strong></td>
<td><strong>11</strong></td>
<td><strong>22</strong></td>
</tr>
</tbody>
</table>
RDMI Functional Framework

- The infrastructure supporting the stages of the RDM lifecycle is built on partnerships with stakeholders and respects the diversity of research data arising from a spectrum of domains.
The Portage Research Data Management Infrastructure (RDMI) Functional Framework specifies the major functional requirements of the stages of the RDM lifecycle. These major functions are divided into subfunctions and aligned with microservices to support the digital services needed by both project level RDM and institutional support for long-term data access and stewardship.
An Example of a Research Data Repository Functional Framework

Microservice Layers

- Data discovery
- Data file transfer: preservation
- Data file sharing: access
- Data access controls
- Data preservation processing
- Data file transfer: deposit

Data Asset Management

Storage
RDM Platforms

- Partner with infrastructure providers to offer research data management (RDM) platforms
  - Launched DMP Assistant in October 2015
    - Introduced customized spaces for institutions in the summer of 2016
    - Finalizing a help desk ticketing service
    - Collaborate in the unification of the codebase for DMP Online (DCC) and DMP Tool (CDL)
    - Completed a pilot with SSHRC and 13 research projects.
DMP Assistant is a bilingual tool for preparing data management plans (DMPs). The tool follows best practices in data stewardship and walks researchers step-by-step through key questions about data management.

**Step 1** Sign up with DMP Assistant

**Step 2** Sign in and select a template under Organizations. The Portage template is the default.

**Step 3** Answer the questions that are relevant to your work. Guidance and examples are provided.

**Step 4** Revisit the tool throughout your research to review or revise your answers.

Please note that we are currently working on single sign-in authentication. For now, please create a new DMP Assistant account. You will have the option to link your DMP Assistant account to your campus ID when that feature becomes available.
Create a new plan

Please select from the following drop-downs so we can determine what questions and guidance should be displayed in your plan.

If you aren’t responding to specific requirements from a funder or an institution, you can choose the Portage Data Stewardship Template. The Portage Data Stewardship Template is based on internationally accepted standards and best practices. It has been prepared and is maintained by a group of research data management experts from research libraries across Canada.

---

To see institutional questions and/or guidance, select your organization.

Portage

---

Choose a template

Portage Template

---

Create plan
• Entered a partnership with Compute Canada in January 2016 to develop an integrated data repository with preservation processing and discovery services.
Federated Research Data Repository
RDM Platforms

● Federated Research Data Repository
  ○ Components of the repository and discovery engine are complete;
  ○ UBC Digital Collections user interface adopted for searching in Dec 2016
  ○ Currently in alpha test mode;
  ○ Beta mode during summer 2017;
  ○ Production in Jan 2018

● Partnerships include Compute Canada, Globus, CARL-Portage, UBC
Partnerships in discussion

● Dataverse North
  ○ Multi-institutional agreement to provide access to Dataverse repository instances
  ○ Scholars Portal middleware development to support archival processing through FRDR

● Jupyter Notebooks and the Pacific Institute of Mathematical Sciences
  ○ Working together to support data repository interoperability with analytic tools provided in Jupyter Notebooks
European Open Science Cloud
Thank you

Chuck Humphrey
chuck.humphrey@ualberta.ca
780-492-9216