



## Portage Progress Report

*For the period of January - March 2019*

---

The following report summarizes Portage activities and accomplishments in the first quarter of 2019 and provides an update on Portage operations and communications undertaken in this period.

Reporting areas include work undertaken towards:

1. Fostering a community of practice for research data management (RDM);
2. Facilitating and providing leadership in the development of RDM infrastructure;
3. Engaging and advocating for RDM with stakeholder communities; and
4. Developing Portage operations and communications.

The mission of Portage is to contribute actively to RDM in Canada through a sustainable network of library-based services and collaborative infrastructure. The following report aims to communicate our progress towards the development of this envisioned national community of practice.

### **Report prepared by:**

Jeffrey Moon (Director, Portage)  
Sarah Wilkinson (Project Officer, CARL/Portage)  
Julie Morin (Project Officer, CARL)  
Lee Wilson (Service Manager, Portage/ACENET)

April 2019

## 1. Fostering a community of practice for RDM

A primary objective of Portage is to build a network of expertise for RDM. Critical aspects of this objective are to coordinate and expand on expertise and services within Canadian academic libraries and to build capacity in specific areas of RDM.

### Activities and Accomplishments

The following update is centered around the work of the Portage Network of Experts.

#### **Preserving Portage White Papers**

- Portage White Papers are being processed for preservation in UBC Open Collections, <https://open.library.ubc.ca/>, an open-access digital repository for research and teaching materials. So far, White Papers by the Data Discovery, Preservation Expert Group, and Training Expert Groups have been published.

#### **Portage Council of Chairs:**

- The Portage Council of Chairs met on February 27, 2019. Chairs heard updates from the Portage Secretariat and provided updates on their progress to date and future work plans. Other points of discussion included Portage funding (CARL member, Tri-Agencies, SSHRC Connection Grants, & ISED), debriefs on the National Data Services Framework Summit, the CANARIE RDM Fund, and a discussion on the current Canadian Repository Landscape.
- The next in-person meeting will be held on May 10th at the University of Toronto.

#### **Curation Expert Group (CEG):**

- The objective of CEG is to identify, evaluate and promote best practices, resources and workflows that support data sharing and reusability for researchers within projects, as well as data within repositories.
- CEG has developed a two-page Data Curation Primer that covers data curation essentials for individuals in curatorial roles. This is available in French and English.
- CEG is in the process of drafting a data curation survival guide, which will be made available online through GitHub and the Portage website.
- CEG continues to collaborate with the Research Intelligence Expert Group (RIEG) to design an information-gathering and outreach strategy/process for identifying data curators, expertise, and activities in Canada as part of a broader 'network-building' initiative.
- CEG continues to work with the FRDR team on curation workflow testing and refinement.
- Portage, in partnership with McMaster University Library, has been awarded a Connection Grant to develop and hold a Canadian Data Curation Forum. This is intended

to help establish a Canadian data curation network (following the model of the US Data Curation Network, or DCN) as well as provide training in collaboration with the Training Expert Group (TEG).

#### **Data Discovery Expert Group (DDEG):**

- DDEG contributes expertise to the Portage Network around the discovery of Canadian research data and the development of a national data discovery service.
- DDEG's FRDR Discovery Service WG continues its work identifying new Canadian data repositories for harvest into the national discovery layer, creating briefing documents to inform both the library and repository community about the FRDR Discovery Service, and mapping harvested subject keywords to the Faceted Application of Subject Terminology (FAST) schema using OpenRefine.
- DDEG has instituted a new working group to undertake the FAST subject mapping project. A call for expressions of interest went out in March 2019, and the group is expected to hold its first meeting in early May 2019.
- In coordination with the CANARIE-funded Geodisy project, DDEG and the FRDR development team are exploring how to integrate geospatial search into FRDR's discovery interface.
- Work continues to keep the Re3data list of [Canadian repositories](#) up-to-date.
- DDEG continues to collaborate with the FRDR Discovery Service WG to identify repositories for addition to the national discovery layer. The full list of collaborating repositories can be found [here](#).
- A FRDR Discovery two-page primer is in process and will be available in April or May 2019.

#### **Data Management Planning Expert Group (DMPEG):**

- DMPEG continues to oversee development of the Portage national data management planning platform, the DMP Assistant. They are gathering DMP 'exemplars' from different disciplines, to assist researchers in drafting DMPs.
- DMPEG and Portage continue to work closely with the University of Alberta on the DMP Assistant 2.0 update/migration. The working plan is still to have DMP Assistant 2.0 launched in 2019.
- Work of the two new DMPEG Working groups continues::
  - The Exemplars WG continues to seek DMPs to add to those already identified.
  - The Policy WG continues to look at DMP vetting and assessment, by looking at the work of other jurisdictions in this area.
- DMPEG is also working on a one-page 'Good Enough' guide to the DMP Assistant.
- The Chair of DMPEG, Carla Graebner, will be on academic leave for 12-16 months starting in June 2019. Two interim co-chairs have been identified: James Doiron, from the

University of Alberta, and Maggie Neilson, from Acadia University. We wish Carla all the best on her leave, and welcome James and Maggie in their new roles.

- A new DMPEG subgroup has been formed: DMP Repositories Working Group. This group will develop a searchable, accessible, and multidisciplinary repository of example data management plans for the research community to consult while writing their own.

#### **Preservation Expert Group (PEG):**

- PEG continues to provide Portage with advice and guidance on RDM infrastructure developments supporting the long-term stewardship of research data, metadata, and code.
- PEG completed development of its current [work plan](#) which provides a high-level overview of proposed priorities and related tasks for the PEG over the next 2 years.
- A subgroup of the PEG is working on a position description for a 'Preservation Coordinator' position, in anticipation of possible funding developments for DRI.

#### **Research Intelligence Expert Group (RIEG):**

- RIEG provides Portage with intelligence on RDM in Canada and are responsible for managing materials and outputs of the Canadian RDM Survey Consortium.
- RDM Roadmap:
  - Taxonomy refinement is underway and will be published soon
  - Roadmap coming to a close - this will set research priorities for RIEG moving forward.
- Format and objective of RIEG will change as the focus will shift from the 'roadmap' project to helping EG/WGs gather evidence and prepare studies/reports on priority topics. To this end, two new working groups have been formed: Current State of RDM Capacity and the Policy/Strategic Planning Evaluation Working group

#### **Training Expert Group (TEG):**

- The primary objective of TEG is the development and maintenance of a high-level 'continuum of training' program to support researchers and RDM specialists across Canada.
- The DMP Assistant training modules are under review by TEG and will be available in Spring 2019. Thanks to OCUL Scholars Portal for providing ongoing access to and hosting of the learning management system that underpins these modules.
- A new Working Group, Repositories 101, has been formed to develop another set of training modules. The group is co-chaired by Vivek Jadon and Laurence Horton.
- Work is ongoing between TEG and other Expert Groups to develop 'Good Enough' one-page quick guides. Some groups have also chosen to develop two-page primers as well. Completed and in-progress outputs and activities include:

Four Good Enough Guides completed: Curation, RDM, FRDR, DVN

Two in progress: DMP, Data Repositories

Four Primers completed: Curation, RDM, FRDR Discovery, DVN (two of these are not currently available but will very shortly)

Other work In progress: DMP Primer, Data Repositories Primer

Two Training Modules completed: RDM 101, CHIR, plus DMP Assistant is in review

Two In progress: Data Repositories 101, Dataverse

In-person training by Portage or its members

- McGill RDM Training Day, Montreal (James Doiron)
- CEGEP RDM Symposium, Quebec (Portage Director)
- University of Alberta RDM Week (James Doiron)

Upcoming activities:

- IASSIST 2019, Sydney, Australia (a number of experts will be doing RDM and Portage-related poster sessions and presentations; Portage Director will be giving a presentation; members of TEG, DMPEG, and RIEG will be presenting and/or giving workshops;
- BCI DMP Training webinars (Danny Letourneau)

Webinars: Portage is working on coordinating webinar-based training within the larger RDM ecosystem, partnering with other stakeholders where appropriate.

- All of these outputs and activities feed into the ‘continuum of training’ model TEG is developing to assist Expert and Working groups with the creation of training materials in support of the larger national community of practice around RDM.

## 2. Facilitating and providing leadership in the development of RDM infrastructure

Portage seeks to advance the development of national platforms for planning, curating, preserving and discovering research data. This will require working with infrastructure providers - both in the library community and elsewhere - to develop new tools where gaps exist and to bridge systems where interoperability is needed.

### **Activities and Accomplishments**

#### **DMP Assistant:**

- No immediate progress has been made on migration of the Portage DMP Assistant to a new merged Digital Curation Centre (DCC) and California Digital Library (DCL) codebase (known as DMP Roadmap). Funding developments may help jump start this project and there is still hope that the migration will take place sometime in 2019. Thanks go to the University of Alberta for continued and generous in-kind support and engagement with the DMP Expert Group and the Portage Secretariat.

#### **Dataverse North Working Group (DVN):**

- DVN was formed to contribute to a community of practice for libraries using or interested in using the Dataverse repository platform for research data in Canada. Work is being advanced by three sub-groups investigating gaps and opportunities that could be addressed by nationally coordinated strategies and services for Dataverse; including, the Business Models Sub-Group, the Training Sub-Group, and the Metadata Sub-Group.
- As of this writing, we understand that a final Service Level Agreement (SLA) between Scholars Portal/U of Toronto and BCI member institutions had been arrived at. This is a big step toward the vision of a national instance of Dataverse (as proposed by the Portage Business Models Sub-Group). Work continues on improving 'internationalization' features of Dataverse (thanks to BCI for their assistance in this). Portage and the Portage Dataverse North Working Group continue to monitor these developments and maintain channels of communication with the stakeholders involved.
- DVN continues to work with the Training Expert Group (TEG) to develop repository-related training materials. The goal is to develop multi-modal materials with priority on introductory manuals, videos, and workshops aimed at librarians and library staff. It has produced a one-page 'Good Enough' guide as well as a two-page primer, both of which are now available.
- The DVN Metadata Subgroup has developed a common 'base metadata template' for Dataverse using standard vocabularies and identifiers. This will be available in Spring 2019. Additionally, this group is working to facilitate the development and adoption of

a common approach to labelling, describing, and organizing data files in Dataverse repositories.

### **Federated Research Data Repository (FRDR):**

- FRDR is a joint project between CARL Portage and Compute Canada to create a national discovery layer for Canadian research data and a scalable repository platform capable of handling large data files easily.
- FRDR is currently operational in a “limited production” capacity, which means that:
  - Anyone may use the service to search for and download data from the FRDR repository or from a growing catalogue of Canadian data repositories.
  - FRDR’s repository services are being offered to a growing list of Canadian researchers. This period saw the [Water Institute](#) at the University of Waterloo join as the latest limited production partner. Ranked among the top 10 water research institutes in the world, the Water Institute is a leader in water research and education. See more on the [Water Institute’s collection page](#).
  - Discussions are underway with several new research groups who could join FRDR as limited production partners in 2019.
- FRDR now offers special collections to large research projects who will be depositing many datasets. Collections allow for branding, project descriptions, links to other online project resources, and collates datasets on a single page. See the [Mountain Legacy Project collection](#) for an example.
- In the fall of 2018, CARL contracted the services of a privacy consultant to review FRDR policies and procedures to ensure that we are in compliance with legislation and best practices. A summary of this work is as follows:
  - Conducted a thorough review and analysis of privacy protocols and documentation related to the Federated Research Data Repository (FRDR)
  - Provided 16 ‘Privacy Administration Recommendations’ for FRDR
  - Drafted a new Privacy Policy
  - Recommended revisions to the Terms of Use to better align with the revised Privacy Policy
  - Provided comments and suggested edits on several FRDR policies
- The FRDR Steering Committee, a group comprising Portage Chairs, members of the Portage Secretariat, and executive members of Compute Canada, approved the FRDR policies drafted by the FRDR Policy Working Group. These included the Terms of Use, Privacy Policy, and several operational policies. The policies were subsequently brought to the CARL Director’s Portage Steering Committee for discussion. The committee raised concerns around legal aspects of the Terms of Use, which are now under review. The policies will be translated and made available on the FRDR website upon approval.

- A priority for Portage in 2019 is moving FRDR out of limited production and into a full service launch.

#### **FRDR Service Model Working Group (FRDR-SM):**

- The FRDR-SM WG was formed to draw on community expertise in the development of service and business models, policies, and user experience and operational workflows for a federated research data repository platform.
- The **Policy Working Group** has completed the development of several policies in support of operating a national-scale data repository service. Drafted policies are currently under review by the CARL Director's Portage Steering Committee. In 2019, the group will work on Service Level Agreements between FRDR and Compute Canada hosting sites, as well as other storage providers.
- The **User Experience and Training Working Group** continues to provide suggestions to improve the FRDR platform user interface design. In conjunction with TEG, the group recently completed a one-page 'Good Enough' guide for FRDR that is available on the Portage website. A call for expressions of interest went out in March 2019 with the aim of gaining several new members to undertake work in its second year mandate where the focus will be on developing more training materials for FRDR, including an online training module.

#### **Repository Landscape Coordination Efforts:**

- Portage has collaborated with Scholars Portal to develop primers and use-case documents for FRDR and Dataverse (*also acknowledging the broader repository landscape in Canada*). These will be translated and made available on the Portage website.
- Portage is creating a new Data Repositories Expert Group (DREG). The Terms of Reference for this group have been drafted, and work is underway to establish a Chair and membership under these terms. As currently envisioned, this group will convene selected representatives from key repository stakeholder communities to provide high-level coordination for platform-specific working groups and to bring a broad and cohesive approach to repository development in Canada. DREG will complement and support the ongoing work of the FRDR and Dataverse North groups, as well as adding consideration of the many domain repositories in Canada.



### 3. Engaging and advocating for RDM with stakeholder communities

Portage seeks to advance the development of national platforms for planning, curating, preserving and discovering research data. This will require working with infrastructure providers - both in the library community and elsewhere - to develop new tools where gaps exist and to bridge systems where interoperability is needed.

#### **Activities and Accomplishments**

##### **Collaboration with Regional Library Councils:**

- The Director continues to meet monthly with representatives of the four Regional Library Consortia. In addition, regional representatives have been invited to add regional updates to Portage Updates released periodically to the community.

##### **Inter-organizational Working Group on Responsible RDM Practices for Sensitive Data:**

- The focus of this working group is to provide practical guidance in managing sensitive data by identifying best practices from both the international community and Canadian experiences.
- Subgroup 1 continues to work on its mandate of developing deposit-friendly text for ethics applications & informed consent, data access agreements, and training materials; The subgroup has developed a Work Plan Grid to identify and prioritize objectives. This grid also includes a section with definitions to ensure common understanding in the group and among consumers of this group's outputs. Thanks to Brenda Gagné and Rachel Zand for continuing to co-chair this subgroup.
- Subgroup 2, looking at issues surrounding indigenous data, has identified a number of action items. The first of these will be to develop a data management planning (DMP) template that addresses indigenous research data issues. These cover the gamut of 'sections' in the DMP Assistant, from data collection through legal and ethical issues.

##### **National Data Science and RDM Training Meeting**

- On January 22, RDC organized a meeting of key RDM stakeholders to discuss coordination of training for RDM and data science initiatives. Included in the meeting were representatives from RDC, CARL Portage, CODATA Canada, World Data Systems, College and Institutes Canada, Centre de documentation collégiale, SSHRC, NSERC, CIHR, and ISED. CARL's Executive Director and the Portage Director and Service Manager participated and spoke to the excellent work being undertaken by the Training Expert Group. Plans are underway to initiate a pilot training group under the TEG that would provide training in direct support of the forthcoming Tri-Agency Data Management Policy.

### **CANARIE RDM Funding Grants**

- On January 23, CANARIE convened a workshop of successful RDM grant holders to discuss progress and explore potential synergies among their projects. The Portage Director and Service Manager attended this workshop. Presentations from the various groups are available on [RDC's Zenodo page](#).

### **National Data Services Framework Summit**

- On Jan 24-25, several members of the CARL Portage team participated in the National Data Services Framework Summit held in Kanata, Ontario. The Portage Director gave a presentation on the work of the Network of Experts and emerging national-scale RDM services, focusing on the national data discovery layer and national repository options. Round table discussions were held on a variety of topics related to the future state of RDM in Canada.

### **ISED Funding update**

- At the NDSF Summit, Sinead Tuite of ISED gave a presentation on the government's plans for DRI funding in Canada which would see the formation of a new organization responsible for Advanced Research Computing, Research Data Management, and Research Software. CARL Portage has met with Compute Canada and RDC to coordinate on planning for the governance and structure of this new organization. At the request of ISED, Portage submitted a 'priority budget' in late 2018 to address medium-term urgent needs to maintain momentum and advance RDM in Canada.
- A high-level Data Management Roadmap document, drafted by a core group of RDM stakeholders (from CUCCIO, IT, RDC, CARL Portage, VPR, Compute Canada, and Ocean Networks Canada), was submitted to ISED on Friday, March 29th. This document is intended to help ISED frame a 'contribution agreement' which is the precursor to flowing funds to the initiatives outlined in the Roadmap.

### **Successful SSHRC Connection Grant**

- A Portage/McMaster application for funding to support a national Curation Workshop was awarded \$18,600. The Forum will be held in October 2019. This event will combine three related goals: (1) providing a community-building stakeholder forum to better understand the current state and unmet needs of data curation practice in Canada, (2) professional development for data curators, and (3) a clear articulation of the next steps for the development of this profession in Canada. This event will play a vital role in developing and teaching the data skills needed to ensure researchers are able to meet the requirements of funders (e.g., the Tri-Agency Research Data Management Draft Policy) and FAIR data requirements in pursuit of objectives such as Open Science, Open Government, and Open Access. This event is part of a long-term, ongoing strategy for the establishment and maintenance of Canadian data standards and practices.

### **CARL Portage & the Montreal Neurological Institute (MNI)**

- On Wednesday, March 27th, the Montreal Neurological Institute hosted an international workshop, 'Making Open Neuroscience Infrastructure Interoperable, 2.0'. The Portage Director was invited to give a keynote presentation on behalf of CARL Portage, highlighting our work in responding to the RDM needs of researchers and the key role of domain specialists in further refining and addressing these needs. The Portage presentation was very well received, and sparked a lot of subsequent conversations. Many in this group were not aware of the emerging Tri-Agency RDM policy, and only a few were aware of the tools, platforms, and services available through Portage. This group was very aware, however, of the crucial importance of metadata standards in their domain, and have taken some good first steps toward addressing this issue.

### **Portage is negotiating an agreement with Clarivate Analytics to make Canadian research data discoverable in their Data Citation Index (DCI):**

- If an agreement is reached, FRDR and its partner repositories will benefit from the provision of DCI's Article-Match-Retrieval system to provide data-use and citation metrics, a free DCI account for each of the 'data providers' feeding metadata into the FRDR discovery layer, and discoverability in an international metadata aggregator.

### **Scholarly Kitchen Blog**

- The CARL Portage Secretariat and a representative of the Scholarly Kitchen Blog met to discuss a 'Guest Post' to this popular scholarly publishing blog.

### **Other Engagements**

- Jeff Moon and Amber Leahey from OCUL Scholars Portal gave a progress report on Portage and Portage-supported CANARIE RDM Funded projects at the Ontario Library Association (OLA) Super Conference on Jan 30, 2019. Immediately preceding this presentation, Jane Fry (Data Librarian @ Carleton U, and Chair of TEG) and Meghan Goodchild (RDM Systems Librarian at Queen's U and OCUL Scholars Portal) gave a talk titled: "Research Data Management: What is it and what's happening in Ontario". Both talks were well-attended and generated interesting questions and discussion.
- Jeff Moon presented on RDM and attended meetings at the University of Alberta, Feb 12.
- Jeff Moon presented and attended meetings at Simon Fraser University's 'Data Love-in' on Valentines Day - Feb 14.
- Jeff Moon presented on RDM and attended meetings at the University of Victoria, Feb 15.
- Jeff Moon presented at the COPPUL Directors Meeting, Feb 28.

- Lee Wilson presented at an RDM stakeholders session at St. Francis Xavier University with Meghan Landry, Liaison Librarian and RDM lead at the Angus L. Macdonald Library, Feb 15.
- Jeff Moon delivered keynote speech at the CEGEP Summit on RDM, Quebec City, March 12.
- Jeff Moon was invited to speak at the Compute Canada Stakeholders Meeting, Ottawa, March 20.
- Jeff Moon delivered keynote speech at “Making Open Neuroscience Infrastructure Interoperable”, Montreal, March 27-28.
- Lee Wilson presented a poster entitled “Canadian Research Data Repository Solutions and Innovations” at the Open Science Conference, Berlin, March 20-21.
- Lee Wilson presented at the Drexel-CODATA FAIR-RRDM Workshop 2019, Philadelphia, PA, March 31-April 1.

#### 4. Developing Portage operations and communications

Portage is a CARL initiative and is currently supported through its membership. The goal is to develop a sustainable network built upon the support of the wider research data management ecosystem.

##### **Activities and Accomplishments**

###### **CARL Directors Portage Steering Committee (CDPSC):**

- The CDPSC met on February 11th, to discuss the Portage Update, Quarterly Progress Report and 2019 work plan. Also on the agenda, were member and external funding, ISED's digital research infrastructure strategy (including governance for the new proposed organization), FRDR and Dataverse North (OCUL's Scholars Portal and Portage), and a debrief on the National Data Services Framework Summit and related meetings. The CDPSC also met on March 26th, to further discuss the Data Management Roadmap to be submitted to ISED, and again on April 4th to review draft FRDR policies and use case scenarios for FRDR and Dataverse.

###### **Portage Advisory Committee (PAC):**

- The Portage Advisory Committee met on March 7<sup>th</sup> and discussed the Portage Progress Report, the National Data Services Framework Summit and related Kanata Declaration, funding for digital research infrastructure, and the current Canadian Repository landscape.