



Portage Progress Report

For the period of October - December 2018

The following report summarizes Portage activities and accomplishments in the fourth quarter of 2018, corresponding to the goals outlined in the Portage Business Plan: 2017 - 2018, 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)

January 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.

Portage Document Migration

- Migration of Portage Expert and Working Group documents to a new Google Team Drive environment is complete. All Expert and Working Group documents have been successfully migrated in coordination with their respective Chairs. Thanks to Portage Service Manager, Lee Wilson (and others at CARL and Portage) for taking the lead on this project, and to EG/WG Chairs and members for their flexibility in adopting this new collaboration platform.

Information Management Policy & Code of Conduct

- As a part of the document migration project, Portage developed a 'lightweight' Information Management Policy and Code of Conduct. These documents outline how Portage-related information should be stored and organized, the responsibilities of Expert and Working Group Chairs around membership management, expectations for the conduct of Portage Network members, and the confidentiality of unpublished work and documents. These documents were reviewed and approved by Portage Chairs during the Nov 2018 Council of Chairs meeting and can be found in the Portage Team Drive [Portage Network folder](#) under 'Policies'.

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 November 2, 2018. Chairs provided updates on their progress to date and future work plans. The group approved a new 'Information Management Policy and Code of Conduct', related in part to the way the new Portage "Team Drive" provides the network with access to Expert and Working Group

documents. Other points of discussion included Portage funding, FRDR, Dataverse, Portage communications/messaging, and governance.

- The next quarterly meeting of the Portage Council of Chairs will be held on February 27th.

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 developed a one-page Data Curation Primer that explains the fundamental elements of data curation. This will be available in French and English in January 2019.
- CEG has developed a two-page Data Curation Primer that covers data curation essentials for individuals in curatorial roles. This will be available in French and English in early 2019.
- 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.
- CEG is working to establish a Canadian data curation network (following the model of the US Data Curation Network, or DCN) through interest groups and a possible national training event in collaboration with the Training Expert Group (TEG). To that end, 3 Canadians attended a DCN Training Workshop in Las Vegas in October 2018. They brought home valuable insights on how to approach a similar national training initiative for Canada.
- The CEG welcomed 3 new members:
 - Guillaume Boutard, UdeM
 - Shahira Khair, UVic
 - Qian Zhang, Waterloo

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. The WG is also developing a brief/primer on the FRDR Discovery Service.

- 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. During this period, several provincial, municipal, and domain-specific repositories were added to the national discovery layer, including the St. Lawrence Global Observatory, the City of Edmonton, and Données Québec. The full list of collaborating repositories can be found [here](#).

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 is working closely with the University of Alberta on the DMP Assistant 2.0 update/migration. Due to a variety of factors, launch of DMP Assistant 2.0, originally planned for the Fall of 2018, is now expected in 2019.
- Work of the two new DMPEG Working groups is progressing:
 - The Exemplars WG has identified and compiled 4 discipline-specific DMPs: Oceans Canada, Historical Canadian Census, Tree frog genetic patterns, and the Womens Print History Project.
 - The Policy WG is working on a survey of international approaches to DMP policies, with a focus on vetting and assessment.
- DMPEG is also working on a one-page ‘Good Enough’ guide to the DMP Assistant.
- DMPEG welcomed two new members:
 - Éve Paquette-Bigras, UdeM
 - Adel Labidi, NRC

Preservation Expert Group (PEG):

- PEG provides Portage with advice on RDM infrastructure developments supporting the long-term stewardship and access of research data and metadata.
- PEG is developing a [work plan](#) which provides a high-level overview of proposed priorities and related tasks for the PEG over the next 2 years.

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.

- RIEG held its first meeting with renewed membership (15 members with representation from RDC and colleges) and revisited its Terms of Reference and work plan for the coming year.
- RDM Roadmap:
 - Taxonomy refinement underway, will be published soon
 - Roadmap coming to a close - this will set research priorities for RIEG moving forward, will be reaching out to other expert groups
 - IASSIST poster proposal submitted
- Format and objective of RIEG will change - focus will shift from the 'roadmap' project to helping EG/WGs gather evidence and prepare studies/reports on priority topics.

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 RDM-101 training modules (4) are now available on the [Portage Training Resources](#) website. DMP training modules are under review and will be coming soon. Thanks to OCUL Scholars Portal for providing ongoing access to and hosting of the learning management system that underpins these modules.
- TEG will be circulating a call in January 2019 for a new Working Group to develop a Repositories 101 training module, which will cover Dataverse and FRDR.
- 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.

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:

- Migration of the Portage DMP Assistant to a new merged Digital Curation Centre (DCC) and California Digital Library (DCL) codebase (known as DMP Roadmap) should take place sometime in 2019, thanks to generous in-kind support of the University of Alberta, and specifically, the work of Weiwei Shi at the University of Alberta, in consultation with the DMP Expert Group.
- Features of the new DMP Roadmap version of the DMP Assistant will include improved machine-actionability, ORCID integration, researcher 'opt-in' for DMP sharing, and APIs to facilitate sharing DMPs.
- Work is underway to renew our agreement with the University of Alberta for continued support of the DMP Assistant. We are grateful to the University of Alberta for their strong in-kind support for this platform, given the important role it plays in the RDM landscape in Canada.

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.
- There has been progress made in the development of a national instance of Dataverse (as proposed by the Business Models Sub-Group). There is continued collaboration between OCUL Scholars Portal and BCI in Quebec to improve 'internationalization' features of Dataverse with the goal of providing Dataverse services to Quebec institutions. Of particular note is OCUL Scholars Portal's successful CANARIE RDM fund application to enhance Dataverse for national repository deployment. 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 will be available in early 2019.
- The DVN Metadata Subgroup continues to work on developing a common 'base metadata template' for Dataverse using standard vocabularies and identifiers. 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. Research projects vary widely in both scope and discipline, and include:
 - The Global Institute for Water Security (University of Saskatchewan)
 - Marine Environmental Research Infrastructure for Data Integration and Application Network (MERIDIAN; Dalhousie University)
 - The Mountain Legacy Project (University of Victoria)
 - Stone Tools, Diet and Sociality at the Dawn of Humanity (University of Calgary)
 - Discussions are currently underway with several new research groups who could join FRDR as limited production partners in 2019.
 - A large number of new datasets were added to FRDR this fall as a part of the Mountain Legacy Project (MLP) collection. The FRDR team worked in conjunction with the University of Victoria libraries, with special thanks to Shahira Khair, to curate the MLP datasets in preparation for publication. More information about the project can be found [here](#).
- FRDR usage report:
 - Published datasets as of Dec. 2018
 - 75 datasets published (4.4 TB)
 - 7 in progress (3.8 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
- The FRDR Steering Committee, a group comprising Portage Chairs, members of the Portage Secretariat, and executive members of Compute Canada, has drafted a Storage Services framework document that proposes how storage allocations for FRDR could work when we launch the service in full production. This document is currently being reviewed for approval by CARL and Compute Canada executives.
- CARL has 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. Part of this work will involve reviewing the policies drafted by the FRDR Policy WG and making recommendations where appropriate.
- 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 Subgroup** 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 FRDR Steering Committee. In early 2019 the group will undergo a mandate refresh to determine the next phase of its work.
- The **Workflows, User Experience, and Training Subgroup** 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 will be made available on the Portage website. The group will put out a call for new members in early 2019 to begin work on a FRDR brief and an online training module.

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.
- Four objective priorities were set: (1) Deposit-friendly text for ethics applications & informed consent; (2) Data access agreements; (3) Training materials; (4) Environmental scan of best practices for Indigenous data. The Working Group recently folded its three subgroups into two:
 - (i) the 'Ethics applications, data access agreements, and Training subgroup.' This will be chaired by Brenda Gagné and Rachel Zand.
 - (ii) the 'indigenous data' subgroup which will continue its exploration of issues surrounding indigenous data with the goal of producing an environmental scan document. It has also set an interim goal of producing smaller guides to key issues relating to Indigenous data, which will inform the larger project.

Other Engagements

- Portage Director and CARL Executive Director attended the CANARIE Summit, October 2-3, 2018, in Ottawa. Titled "Automation Nation", this summit provided attendees with the opportunity to learn about emerging trends in artificial intelligence, augmented reality, and more. RDM was mentioned in some of the talks, and was certainly the topic of many side discussions.
- Portage Director and CARL Executive Director attended a meeting on the future of RDM in Canada convened by CANARIE, October 5, 2018, in Ottawa. This meeting, attended by contributors to the August 2017 LCDRI Data Management report to ISED, provided an opportunity for discussion of the need for close RDM stakeholder collaboration

during the transitional period to the new organization for the national layer of Digital Research Infrastructure.

- Portage Service Manager provided a Portage update to the CNC CODATA meeting held on October 16th.
- Portage Service Manager was selected to attend the Specialized Data Curation Workshop organized by the Data Curation Network from Oct. 17 - 18 co-located with the DLF Forum and NDSA's Digital Preservation 2018 in Las Vegas, Nevada. This 1.5 day long training session brought together library data specialists and discipline and functional experts in a peer-to-peer learning environment for specialized data curation. Shahira Khair and Sandra Sawchuk of the Curation Expert Group were also selected to attend. Information and approaches gathered at this workshop will inform possible workshop offerings in Canada.
- Portage Director served on a panel (with a representative from NSERC) at a Colleges and Institutes Canada 'Applied Research Symposium' held in Ottawa on November 6th, 2018.
- Portage provided an update on activities and progress to CARL Directors at the CARL Fall General Meeting (Nov 7-8, 2018) in Montreal.
- Portage Director attended the CARL-sponsored @ Risk North II: Digital Collections meeting held on Nov 9, 2018 in Montreal.
- Portage Director contributed to an RDM Workshop at the University of New Brunswick on Nov 22, 2018.
- Portage Director and CARL Executive Director attended the Queen's Principal's Symposium: Imagining Our Digital Future in Kingston, Nov 25, 2018.
- Portage Service Manager presented at the Stakeholder's Session of the Compute Canada Federation's Board Meeting held in Montreal, QC on Dec 4, 2018.
- Portage Director, CARL Executive Director, and President of CARL attended the G7 Research Summit-sponsored 'LAC Memory Institutions in the Digital Age' on Dec 5, 2018, in Ottawa.
- Portage Director participated in a panel discussion on the evolution of data access at the Statistics Canada@100 Conference in Ottawa, Dec 6, 2018.
- Portage Service Manager presented at Research Atlantic 2018 held at Mount Saint Vincent University (Halifax, NS) on Dec 5-6, 2018.
- Portage Director presented at a Colleges, Institutes, and Polytechnics VPRs RDM Day on December 7th, 2018 at Centennial College in Scarborough.

- Portage Director contributed to an RDM Workshop at McMaster University on Dec 12, 2018

National Data Services Framework Summit

Portage has been involved with the planning committee for the upcoming RDC-sponsored NDSF Summit (Jan 24-25, 2019). A discussion document will be released in advance of the Summit.

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

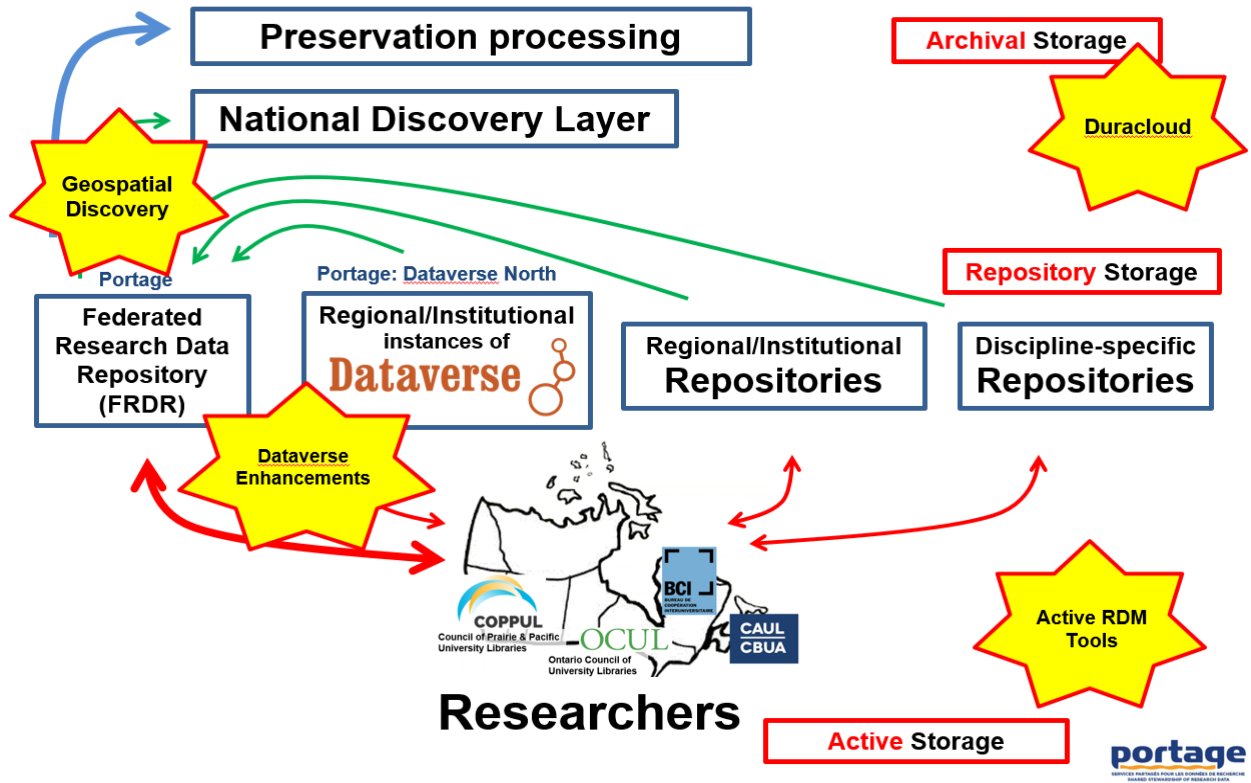
- If an agreement is reached, FRDR and its partner repositories will benefit from the provision of DCI's Article-Match-Retrieval system to provide data-use and citation metrics, a free DCI account for each of the 'data providers' feeding metadata into the FRDR discovery layer, and discoverability in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

CANARIE RDM Funding Call:

Portage provided letters of support for four proposals for the CANARIE RDM Funding call. In the end, all four applications were successful in obtaining funding. Funded proposals supported by Portage include:

- 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.
- The Radiam project will assist research groups in managing their 'active' research data: beginning at the start of the data life cycle up to publication by providing content indexing and searching tools, both stand-alone and also integrated into existing platforms such as HubZero and Open Science Framework. Radiam is a joint project between the University of Saskatchewan and Simon Fraser University, with support from Compute Canada and CARL Portage.
- 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.

An illustration of where each of these projects fits into the broader Portage vision for a Federated RDM Architecture is shown in the figure below.



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

Portage Funding

- Renewed and generous additional financial support from CARL Library Directors, further augmented by strategic funding requested from the Tri-Agencies, has put Portage on firm footing for 2019. At the request of ISED, Portage is working on a 'priority budget' to address medium-term urgent needs to maintain momentum and advance RDM in Canada.

CARL Directors Portage Steering Committee (CDPSC):

- The CDPSC held a meeting on October 23rd, in order to plan for the Portage briefing at the CARL Fall General Meeting on November 7th.

Portage Advisory Committee (PAC):

- The Portage Advisory Committee met on October 24th and discussed the Portage Progress Report, submissions to the CANARIE RDM Fund, responses to the Tri-Agency Consultation on their draft Research Data Management Policy, ISED's digital research infrastructure strategy, and 2019 CARL member funding.