



Portage Progress Report

For the period of July 2018 - September 2018

The following report summarizes Portage activities and accomplishments in the third 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)

October 2018

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 centred around the work of the Portage Network of Experts.

Portage Document Migration

- Portage is migrating Expert and Working Group documents from the current Google Drive environment to a Google 'Team Drive' account that will ensure continuity of, and access to, the work of these groups over time. Most Expert and Working Group documents have been migrated in coordination with their respective Chairs.

Information Management Policy & Code of Conduct

- As a part of the document migration project, Portage is developing a lightweight Information Management Policy and Code of Conduct. These documents will outline how Portage-related information should be stored and organized, the responsibilities of Expert and Working Group Chairs around membership management, and expectations for the conduct of Portage Network members and the confidentiality of unpublished work and documents. The Policy and Code of Conduct will be reviewed by Portage Chairs during the Nov 2018 Council of Chairs meeting.

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 next quarterly meeting of the Portage Council of Chairs is scheduled for November 2nd, 2018. In addition to hearing updates from the Portage Secretariat and each of the Expert Group Chairs, agenda items will include reviewing a draft information management policy and code of conduct agreement, and coordination among groups.

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 is developing a Data Curation Primer, a one-pager that explains the fundamental elements of data curation.
- CEG is developing a Data Curation Survival Guide, which covers data curation essentials for individuals in curatorial roles.
- CEG is collaborating 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.
- CEG is in the process of drafting a data curation white paper on developing a national approach to data curation in Canada.
- 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).

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.
- Based on the work of the FRDR Discovery Service WG, several new data repositories have been added to the national discovery layer, including the [Hakai Institute](#), the [St. Lawrence Global Observatory \(SLGO\)](#), and several provincial and municipal data repositories. The full list of collaborating repositories can be found [here](#).

Data Management Planning Expert Group (DMPEG):

- DMPEG continues their work overseeing the 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.

- DMPEG has two new Working groups:
 - Exemplars WG: identifying and compiling discipline-specific DMPs.
 - Policy WG: looking at criteria for evaluating DMPs, e.g. for adjudication purposes.

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 planning its next steps based on its recent White Paper, [Research Data Preservation in Canada](#), released last spring.

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 put out a Call for Participation on September 6, 2018. 10 new members will be added in the coming months.

Training Expert Group (TEG):

- The primary objective of TEG is the development and maintenance of a high-level program for supporting RDM training across Canada.
- The RDM-101 and DMP online training modules are under development in Scholars Portal, which has provided ongoing access to a learning management system.
- TEG will be circulating a call for a new Working Group to develop a Repositories 101 training module, which will cover Dataverse and FRDR.
- TEG is working with other Expert Groups to develop additional 'Busy Bee' one-page quick guides (e.g. FRDR and DVN plan to add to this series).

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:

- Planning for migration of the Portage DMP Assistant to a new merged Digital Curation Centre (DCC) and California Digital Library (DCL) codebase (known as DMP Roadmap) is underway, thanks to the in-kind support of the University of Alberta, and specifically, the work of Weiwei Shi at U of Alberta, in consultation with the DMP Expert Group.
- Version 2.0 of the DMP Assistant promises exciting new features, including a foundation for machine-actionable DMPs, ORCID integration, researcher 'opt-in' for DMP sharing, and APIs to facilitate sharing DMPs. Launch date to be determined, based on progress made.

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.
- Work toward a national instance of Dataverse (as proposed by the Business Models Sub-Group) is evolving. OCUL-Scholars Portal is working with BCI in Quebec to improve 'internationalization' features of Dataverse with the goal of providing Dataverse services to Quebec institutions. Portage and the Portage Dataverse North Working Group are monitoring these developments and maintaining channels of communication with the stakeholders involved.
- DVN is working 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.
- The DVN Metadata Subgroup is continuing its 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)
 - Diversity Cape Breton (Cape Breton University)
 - The Water Institute (University of Waterloo)
 - Ocean Networks Canada-DFO (University of Victoria)
 - As well as individual researchers from several universities
 - Discussions are currently underway with several new research groups who could join FRDR as limited production partners in the fall.
 - Several new datasets were added to FRDR over the summer, and we expect submissions to increase during the fall with the release of the Mountain Legacy Project collection. The FRDR team has been working in conjunction with local RDM contacts to curate datasets in preparation for publication.
- The FRDR Development Team is testing browser upload capabilities to allow users to upload small datasets without requiring Globus File Transfer. This addition will make it easier for researchers to make use of FRDR. Browser upload will be added to FRDR in fall 2018. For more details on the ongoing development of FRDR, see the [Change Log](#).
- 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 undertaken the development of several policies in support of operating a national-scale data repository service. These efforts are nearly complete, with the majority of the policies now drafted and ready for review by the FRDR Steering Committee.
- The **Workflows, User Experience, and Training Subgroup** continues to provide suggestions to improve the FRDR platform user interface design, with feedback on the new browser upload UI being collected. During the spring/summer of 2018, this group conducted an extensive overhaul of all user-facing documentation associated with the service. The new documentation is now available in English on the FRDR website, and will be translated into French shortly.

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

Portage Response to Tri-Agency RDM Policy

- The Portage Secretariat submitted a final response to the Tri-Agency RDM Policy at the end of August, following a period of open consultation starting in June, and culminating in three Town Hall stakeholder meetings held in July. The Tri-Agencies expressed gratitude for the work done by Portage in gathering, synthesizing, and reporting this feedback.

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 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 combined two subgroups to create the 'Ethics applications, data access agreements, and Training subgroup.' This will be chaired by Brenda Gagné and Rachel Zand. The second subgroup is continuing its exploration of issues surrounding indigenous data with the goal of producing an environmental scan document.

Other Engagements

- Jeff Moon (Director, Portage) traveled to Lakehead University on September 19th to lead a Research Data Management (RDM) Workshop with 74 in-person registrants and 21 online registrants.
- Lee Wilson (Service Manager, Portage) traveled to São Paulo, Brazil to give a talk on Portage and RDM activities in Canada, as well as meet with researchers, librarians, and research administrators involved in developing RDM services both at the

local/institutional and state levels (Sept. 18-21). A representative from DANS (Data Archiving and Networked Services) -- the Netherlands Institute for permanent access to digital research resources -- also delivered a presentation about DANS' services and RDM developments in the Netherlands and the European Union. Travel was sponsored by the São Paulo Research Foundation (FAPESP) -- a public foundation with the aim of providing grants, funds and programs to support research, education and innovation of private and public institutions and companies in the state of São Paulo.

- Jeff Moon (Director, Portage) participated in a meeting (replacement for G7 Research Summit, cancelled due to tornadoes in the Gatineau area) on September 24th, 2018. This was a working session hosted by RSC President, Chad Gaffield and International Secretary, Marie D'lorio in collaboration with Canada's Chief Statistician, Anil Arora and colleagues. This morning meeting was held in lieu of the planned Summit.
- RDC hosted a meeting of the authors of the LCDRI DM (Data Management) Report to ISED (submitted to ISED August 31, 2017). The group discussed a number of RDM-related issues, and made progress toward drafting an 'environmental scan' of RDM in Canada based on the LCDRI DM document.
- Portage and Compute Canada held a meeting on September 5th, 2018 to discuss, preliminarily, the intersection of work being done by Portage and ARC under the discussion paper distributed by ISED; further discussions, involving other RDM partners (e.g. RDC) will be necessary going forward.

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. The four applications in contention for RDM funding (successful projects not yet announced by CANARIE as of this writing) were:

- 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.

- Improved integration with 'Active' RDM tools through metadata aggregation and enhancement plug ins/APIs (for example, HubZero and Open Science Framework). Proposal to be led by the Compute Canada RDM team at USask (the same team that developed FRDR).
- 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 announcement of successful projects is anticipated in late October/early November, 2018.

Discussions with Erudit and PKP

- Jeff Moon (Director, Portage) participated in discussions with [Erudit](#) and [PKP](#) regarding treating their corpus of academic journal articles as 'data' for purposes such as text mining. This also included discussions about the continuous flow of articles and the challenge this poses when it comes to defining a 'dataset' for citation and/or preservation purposes. Further discussions will take place. Ultimately, Portage provided a letter of support for Erudit in their application to the CANARIE RDM Funding call.

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 held a one-hour teleconference on September 21st, where they discussed submissions to the CANARIE RDM Fund, shared use of Scholars Portal Dataverse, responses to the Tri-Agency Consultation on their draft Research Data Management Policy, and ISED's digital research infrastructure strategy. They also discussed member funding models for Portage in 2019 and scheduled an additional meeting for October 23rd, in order to plan for the Portage briefing at the CARL Fall General Meeting on November 7th.

Portage Advisory Committee (PAC):

- The next regularly scheduled meeting of the PAC will take place on October 24th. Agenda items will include the Portage Progress Report, submissions to the CANARIE RDM Fund, responses to the Tri-Agency Consultation on their draft Research Data Management Policy, and ISED's digital research infrastructure strategy.
- In September, an article on Portage, by Steven N. Liss, Vice-President (Research and Innovation) at Ryerson University and Chair of the Portage Advisory Committee, was published in University Affairs. The article, titled 'Addressing Gaps in Canadian Research Data Management. A comprehensive overview of the Portage Network', describes how the Portage Network of Experts and other stakeholders are working to develop the tools, resources, and services Canadian researchers need to meet the RDM challenges ahead.

Portage Project Officer:

- CARL hired a new Project Officer, Sarah Wilkinson, who started working with CARL-Portage on September 4th, 2018.