

Portage: Supporting Canadian innovation through shared expertise and stewardship of research data

Brief for CARL Directors, November 2014

In March 2014, CARL launched Project ARC to develop a library-based research data management (RDM) network in Canada. The proposed name for the network is Portage: supporting Canadian innovation through shared expertise and stewardship of research data” (en français, “Portage: Soutenir l’innovation canadienne par le partage d’expertise et la gestion des données de recherché”).

The Portage network will have two major components:

- A distributed centre of expertise for research data management, and
- A national preservation system for research data that will evolve and expand over time

Network Centre of Expertise

Since its launch, Project ARC has collected a comprehensive set of resources to support data management planning. These resources were derived from a number of sites in Canada and internationally and point users to the most up-to-date, relevant and trusted sources about research data management. They include “how to” guides, case studies, training materials and links to other important resources. For the moment, the resources are available on the Project ARC website, but they will eventually be migrated to a permanent national website. In addition, Portage will host a national DMP automated tool to assist Canadian researchers in developing data management plans.

The centre of expertise will also provide consulting services drawing on the expertise of librarians and others from across the country. Specific areas of know-how will include data curation, training, DMPs, discovery, preservation, and privacy-security-ethics. Although the details of the delivery model are still being considered, it is proposed that consulting services will be provided via a national pool of experts from member institutions, who will devote a certain portion of their time to the centre. Much of this will be in-kind support flowing from existing job responsibilities, but there may also be opportunities for members to host experts at their institutions, funded through a central resource pool. The intent is to build human capacity across the country.

National preservation infrastructure for research data

At the same time, Project ARC has also been working on a pilot project to connect the various infrastructure and service components needed for a national preservation network. This intersects directly with the centre of expertise concept: advice and support for researchers should be accompanied by viable technical solutions. The pilot is being undertaken in close collaboration with Compute Canada, Research Data Canada, and some of the domain data centres (Canadian Astronomy Data Centre and C-Brain) to ensure that it will be both inclusive and interoperable, two important characteristics to be eligible for external funding.

To date, repository software has been installed at two different Compute Canada sites. Project participants at SFU are connecting Islandora repository software to Archivematica (a preservation platform), which have been installed on top of the Compute Canada stack at SFU, and Scholars Portal has been working with Compute Canada to connect Dataverse with Archivematica, and install them on CC resources at the University of Toronto. The aim of these activities is to improve data flow so that data can easily move from data repositories into an environment where preservation and replication can be undertaken. The image attached at the end of this document illustrates the vision for the infrastructure component of the network.

The pilot will soon begin to ingest data from selected researchers of different disciplines at the two different locations. This will identify any issues with the systems and workflows. Once any problems have been addressed and the workflow has been stabilized, the network will expand to include other repositories. The ultimate aim is to enable all interested academic libraries to participate, whether or not they have their own local infrastructure, by coordinating shared repositories and services under a cost model that recognizes varying institutional investments and needs.

Governance

There will be a need for a central secretariat to coordinate activities of the network, manage finances and provide institutional support. The network will be governed by participating institutions, with a formal mechanism for other stakeholders to provide input and advice. Mindful that the creation of a new stand-alone organization would lead to unnecessary administrative costs and complexity, a number of governance models are being developed for consideration by the wider community. We welcome thoughts and suggestions from CARL members.

Funding

The Portage network will require approximately \$1 million CAD per year in order to administer the centre of expertise and further develop and manage the infrastructure. It is proposed that the network be funded for a period of three years, during which there will be ongoing assessment of its success and further delineation of future needs. To date, a number of possible funding models have been identified, for discussion with CARL Directors.

In the coming months, Project ARC will continue to consult with CARL members and other stakeholders to identify preferred models for funding and governance. Community input will be critical to ensure initial buy-in and ongoing support for the network.

How can your institution participate in Portage?

Right now

- Contribute to the development of information resources and national guidance documents

- Provide input to the Project ARC Working Group, particularly in the development of the governance and funding models

In the future

- Contribute knowledgeable staff to the pool of expertise
- Connect your data repository to the preservation network
- Participate as a user of the services and a stakeholder in their development

The Project ARC Working Group has members from all four regional academic library associations (CAUL, COPPUL, OCUL and Quebec) as well as CRKN, CARL, and some of Canada's top research data management experts.

For more information, or to get involved, please contact Kathleen Shearer, coordinator of Project ARC or Martha Whitehead, chair of Project ARC Working Group.

