MULTI-STAKEHOLDER ENGAGEMENT IN RESEARCH DATA MANAGEMENT

Martha Whitehead
Vice-Provost and University Librarian, Queen's University, Canada
martha.whitehead@queensu.ca

Donna Bourne-Tyson
University Librarian, Dalhousie University, Canada
donna.bourne-tyson@dal.ca

Abstract

Research data management (RDM) services have become a high priority for government agencies and post-secondary institutions across Canada in recent years. There is a strong sense of urgency: Canada has lacked coherent national strategies such as those in Australia or the UK, and at the same time there are growing expectations for sound RDM practices. For example, Canada's Action Plan on Open Government includes deliverables aimed at improving access to publications and data resulting from federally funded scientific activities, and the federal research granting councils will be adopting a *Statement of Principles on Digital Data Management* in 2016 that will be reviewed and revised through continual stakeholder engagement. Amongst those stakeholders are the research universities of the country, including their libraries.

Canadian university libraries have a long history of the kinds of collaborations required in the multi-stakeholder RDM environment, deep experience in developing programs to advance research, and critical expertise in preservation. In 2015, the Canadian Association of Research Libraries (CARL) launched a national research data management network, named Portage, to assist researchers and other RDM stakeholders through a university library-based network of expertise on RDM and through working with multiple stakeholders to develop national platforms for planning, preserving, and discovering research data. This work has proceeded in concert with the RDM priorities of university presidents and research administrators, as well as federal research granting councils and agencies responsible for advanced research computing and the national high-speed optical research network.

This paper will describe approaches taken to aligning RDM services across multiple stakeholder groups in Canada and discuss various factors to be considered in such collaborations.

Keywords: research data management, RDM, collaboration, national strategies, research libraries

The Canadian context

The current Canadian government was elected on a platform that included a promise to make government science fully available to the public. This bodes well for the future of Canada's Action Plan on Open Government, which includes deliverables aimed at improving access to publications and data resulting from federally funded scientific activities [Open Government Canada, 2014]. It also bodes well for research activities across the public sector. The challenge, however, is that Canada has been lacking in national strategies for developing the digital infrastructure required to move from high-level philosophy to practical support for open science. Different aspects of that infrastructure have traditionally operated separately and not as parts of a cohesive whole – at the national and institutional levels – until now.

The Government of Canada provides funding for a variety of entities that support research in different ways. There are three main federal research granting councils, known as the Tri-Agencies, which come together for purposes such as joint policy development. The Canada Foundation for Innovation is another government agency, providing research infrastructure funding. Further government funding supports a national high-speed optical research network (CANARIE) and advanced research computing (Compute Canada), both of which work

collaboratively with regional counterparts. Compute Canada has four regional affiliates -- ACENET, Calcul Québec, Compute Ontario, and WestGrid -- which operate somewhat independently of Compute Canada, but very closely with the universities in the region. Canada is a federal system of provinces and territories, and those regional governments also fund some elements of research and most aspects of higher education.

Existing federal digital infrastructure tends to focus on supporting particular research projects, rather than horizontal services crossing disciplines. Within this distributed landscape, there have been some critical gaps, including vulnerability of the actual research data that is used on the research networks and advanced computing platforms. The emphasis until recently has been on building a robust digital network and providing individual researchers or national research projects with short-term High Performance Computing (HPC) and limited data storage for one to three years; there was no focus on long-term preservation or re-use of research data. Recognizing this issue, many stakeholders across the research landscape came together under the leadership of the National Research Council in 2008 to form a Research Data Strategy Working Group, which held a National Data Summit in 2011. From there, an entity called Research Data Canada (RDC) was formed, with a mandate to work at a policy level with stakeholders across the country and to facilitate the development of on the ground services.

Following on the heels of the 2011 Summit, and broadening beyond research data matters to the full digital research infrastructure, a confederation of agencies and associations including RDC joined together in 2012 to form the Leadership Council for Digital Infrastructure (LCDI) to better coordinate Canadian efforts. LCDI is a stakeholder-driven group that provides a forum for system-wide discussion and collaboration on issues of common interest. The Canadian University Council of Chief Information Officers (CUCCIO) has provided essential leadership support to the group and members include RDC, CUCCIO, CANARIE, Compute Canada, Canadian Association of Research Libraries (CARL), the Canadian Research Knowledge Network (CRKN), CASRAI, members of the U15 Vice-Presidents Research (VPRs) group, and representatives from the federal research granting councils. LCDI hosted two successful summits on digital research infrastructure, in 2012 and 2014. 'Summit 2014' resulted in a number of important discussions among users, research funders, and service providers within the Canadian digital research infrastructure community, as well as significant engagement with federal officials. In addition, it produced a number of important documents that can be found at http://digitalleadership.ca/di-summit-2014/.

Awareness grew during the period of these summits that there are three complementary aspects to a thriving digital infrastructure: the network required to transmit and share the data; the computing power and software tools required to create and manipulate the data; and the expertise and systems required to preserve and re-use the data. These three aspects were managed by one or more of the LCDI partners, and there was no well-defined comprehensive view of interdependency of the three aspects and the various partners' roles at the outset. In 2015, LCDI's response to an Industry Canada consultation noted that like any ecosystem, the national ecosystem for digital research infrastructure is inherently complex with many stakeholders of widely varying capability, capacity, and responsibility [Leadership Council for Digital Infrastructure, 2015].

Within this digital infrastructure landscape, CARL has long been actively involved in data management and was a founding member of RDC. In late 2013, CARL convened a meeting of individuals representing regional groups and institutions that had begun to develop local library-based research data management (RDM) services, with a view to considering what could be achieved collaboratively on a national scale. The aim was to leverage existing strengths, build capacity and fill gaps. In March 2014 a one-year pilot project was launched to develop the concept of a network of expertise and an infrastructure model for data preservation workflows. The infrastructure aspect of the pilot involved a collaboration with Compute Canada, under the umbrella of a "federated pilot" that RDC helped to facilitate.

The pilot project provided evidence of what could be accomplished at the national level through bottom-up collaborations and lightweight coordination, and there was a strong desire to build on this momentum. It was also recognized that a robust network would require significant in-kind contributions from participating partners as well as funding for projects and ongoing operations. CARL committed to hiring a network director for a two-year period, during which the basic services of the network would be established along with a governance structure and funding model. In September of 2015, Chuck Humphrey was seconded from the University of

Alberta to the role of Director of the Portage Network. The network is described in detail later in this paper.

The basic premise of Portage is that a national network will help enable institutions to build their capacity for RDM services, and thus increase the impact of the research being conducted across the country and around the world. For individual researchers, the current gaps may be apparent in their needs for advanced research computing, data storage or services to manage and preserve their data. Those researchers often turn to units at their universities for help – research services offices, information technology (IT) staff, and the information services of the library. These units have always worked together in a variety of ways, and at many institutions they are coming together locally and nationally to address RDM needs.

Institutional Needs

There is a plethora of tools and standards available to support every stage of the research life cycle, but stewardship has been lacking. As Andrew Sallans noted in his keynote at the International Digital Curation Conference 2016, openness and sharing are central to research. He observed that data is lost in many ways throughout the lifecycle, through paywalls, broken links, and lack of discoverability options, and he recommended five percent of research budgets be reserved for data stewardship, and the requirement within a decade for half a million data experts in Europe alone. [Sallans, 2016]

Ideally a university researcher will be able to seek the assistance of a librarian throughout the research life cycle, beginning with the development of a research data management plan (DMP). Research libraries in Canada are building relationships with research services offices and researchers are often referred to a librarian when writing a grant proposal for a new research project. Many Canadian universities are adopting the Portage DMP Assistant, which is based on the Digital Curation Centre's DMPOnline (see below). The creation of metadata is a critical part of the RDM process, to provide context for research data and for it to be discovered and reused later. Institutions are looking to librarians for expert advice in this area, once the importance of metadata creation is understood. Researchers already use library resources, discovery tools and citation management software during the discovery and literature review phase of their research, and often work with librarians for knowledge synthesis or systematic reviews. Using a DMP tool is the next step in the relationship, though an earlier stage in the research process.

Institutional support for the active middle stages of a research project varies, in terms of the provision of a virtual research environment (VRE), HPC, or even a basic service advising on best practices for cloud-based storage. In many situations, the universities leave the researchers to work with Compute Canada and the regional affiliates at this stage. There is awareness of a growing need for digital research infrastructure but universities, provincial, and federal governments are not yet aligned in sharing responsibility for the infrastructure. In comparison to systems such as Australia's Nectar Cloud, Canada is still formulating a coordinated plan for research infrastructure requirements beyond HPC, through the explorations of Compute Canada and the regional affiliates, CARL/Portage, RDC, and the LCDI. At the May 2016 ACENET Annual General Meeting, for example, researchers noted the need for petabytes of storage. There is also a huge requirement for assistance with setup; although large projects are often well served, assistance is needed with small and medium sized endeavors, even in fields such as computer science, biology, and engineering. In the absence of sufficient advice and RDM related services, individual researchers' funding is being spent on very small systems, which are not interoperable, scalable, or the best use of funds. There are implications for libraries, IT departments and for Compute Canada and the regional affiliates about staffing required to support a wider array of services and system configurations.

Institutional support for the latter stages of dissemination and preservation of research data has been considered more the domain of the university libraries than earlier phases, though sometimes in collaboration with IT and/or the research services office. Take-up rates for repository deposits and preservation services have been relatively low, however there is now a requirement that research funded by the three granting councils in Canada be either published in an OA journal or deposited in a repository within a year of publication. The implications of the Tri-Agency Open Access Policy on Publications and soon for data are raising interest levels outside of the libraries, and also expectations.

Capacity to scale up current operations to meet demand is a concern for most university libraries in Canada, and the importance of coordinating efforts through CARL's Portage initiative is well recognized. Issues related to discoverability, citation management, bibliometrics, storage capacity, preservation standards, privacy, ethics, and sustainability abound; international standards and best practices have been gathered and tailored to suit institutional needs, until recently without a strong national voice guiding the coordinated adoption of standards and the development of services.

National Principles

There has been a general sense amongst RDM stakeholders that national and local policy frameworks will be a key factor in the successful development of coordinated, interoperable services. This aspect of RDM has been advanced recently by two initiatives:

- TheTri-Agencies in 2015 invited targeted stakeholders to review and provide feedback on a draft *Statement of Principles on Digital Data Management.* This followed a previous consultation period on a potential policy around the stewardship of research data through their paper *Capitalizing on Big Data: Towards a Policy Framework for Advancing Digital Scholarship in Canada*, to which many stakeholders including CARL responded [CARL, 2013]. In 2016, the Tri-Agencies plan to adopt the statement of principles as a living document to be reviewed and revised as required through continual stakeholder engagement and as the RDM environment evolves in Canada and globally.
- VPRs have been in attendance at the RDC and LCDI summits and subsequent events such as the Strategic Workshop on Research Data Management for Vice-Presidents Research organized by RDC and the University of Alberta. The latter meeting resulted in the document *Research Data Management in Canadian Universities: A Statement of Principles*, released in March 2016. This Statement was adopted by the U15 at the April 2016 meeting.

These guidelines provide strong frameworks for individual institutions as they develop closer working relationships among the IT department, the library, and the research services office, while also determining what is best managed at a regional or national level. VPRs, University Librarians, Chief Information Officers, association Executive Directors, and granting agency Presidents are all becoming better acquainted and trying to coordinate efforts. Emblematic of the urgent need for coordination at both the institutional and national level was the fact that upon arriving at the LCDI Summit 2014, several universities found they had sent between three and six representatives, with various institutional and national responsibilities in IT, libraries, and the research services office, and serving on various national boards and granting councils. While it is clear that changes which are supportive of the developing research data culture are happening in local and national institutions, there remains a need to continue addressing data culture changes with researchers in some domains, where data sharing has not been a tradition. Researchers in a few distinct disciplines do not see a need to prepare data for others to use; in this regard, libraries find themselves in the role of change agents, due to changes in technology as it relates to publishing, content, instruction and research.

Portage Network

The overarching vision for Portage is a future in which Canada capitalizes on the trend towards data-intensive research to become a world leader in research and innovation. Those involved in launching Portage believe that a key element of achieving that future is comprehensive support for research data management at a national scale. Full information about Portage and its services is available at https://portagenetwork.ca/.

At the conclusion of the pilot project that preceded Portage, it was envisioned that an ongoing network would have two interrelated elements to assist researchers and other research data management stakeholders: (1) a library-based network of expertise on research data management and (2) national platforms for planning, preserving, and discovering research data. The Portage Network Framework diagram below provides a view of the network's mandates using a research data life cycle model [Humphrey, 2016].

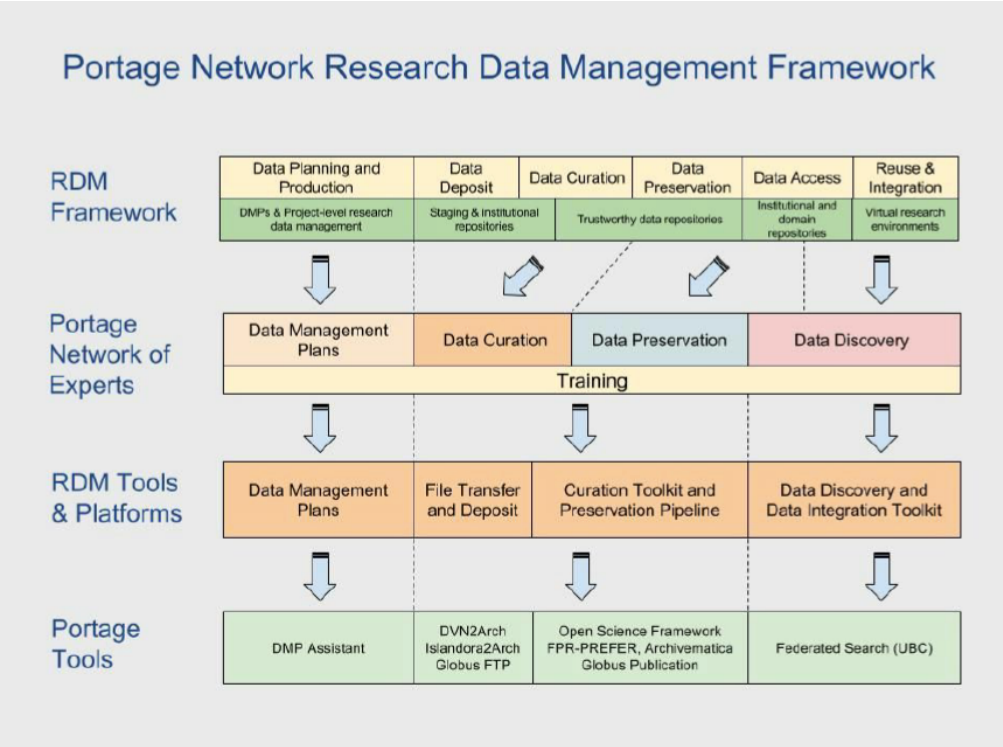

Figure 2: Portage Network Research Data Management Framework

Network of expertise on research data management

The aim of this aspect of the network is to capitalize on expertise and services within Canadian academic libraries and build capacity in specific areas of research data management. A suite of network services, tools, and information resources will be developed and supported by a set of "expert groups" consisting of individuals across the country. These groups rely on the in-kind contributions of the participating individuals' institutions; as well, the Director is in the process of seeking project funding for some aspects of their activities. The expert groups established to date are:

- *Data Management Planning*: This group has launched a bilingual data management planning tool, DMP Assistant (see Network Platforms below). The group has also developed a localization kit for institutions to administer individual library spaces on DMP Assistant.
- *Data Curation*: This group will focus on the tool kit, services, and skills for managing data during research projects, with the intent to facilitate the subsequent deposit of data for preservation and sharing, and for performing longer- term data stewardship.
- *Data Preservation*: Serving as an oversight group for preservation platforms, this group will collaborate with partnerships in developing different preservation components and platforms. It is currently working with Compute Canada to develop a production-level pipeline (see Network Platforms below) and has also taken initial steps to work with middleware projects bridging repositories and the preservation pipeline project (for example, working with Dataverse).
- *Data Discovery*: This group's work relates to metadata standards supporting discovery, and discovery systems for research data across multiple data repositories. It has begun preparing a white paper on current best practices and is exploring the application of the University of British Columbia's federated digital collections discovery system to research data.
- *Training*: The need for training cuts across all aspects of research data management services. Over time it may take the form of online videos, workshops and onsite institutional engagements, as well as information drawn from best practices around the world. The UK Digital Curation Centre has been supportive of Portage from the outset and provides excellent experience to draw upon.

- *Topical groups*: It is anticipated that with sufficient resources the network could address other topics. One of the most pressing is work with the research ethics community to develop best practices for the ethical treatment of data within research projects and its subsequent stewardship at the end of a project's life. There has also been interest expressed in developing cost-effective ways of certifying repositories through a peer review process.

Network platforms for planning, preserving, and discovering research data
The aim of this aspect of the network is to connect the various infrastructure and service components needed for planning, preserving and discovering research data by coordinating infrastructure across the country, filling gaps where tools are missing, and bridging systems where interoperability is needed.

This is a distributed infrastructure model, with an expectation that funding will come from multiple sources and be applied to multiple nodes managed by different stakeholder groups. The ultimate aim is to enable all interested universities to participate, whether or not they have their own local infrastructure. The envisioned platforms provide data management planning tools, and ingest and preservation services that include networked replication storage services, and will be based on standards that ensure interoperability across nodes and data types. In addition to ingest and preservation, a complementary set of services will support the discovery of data contained in data repositories across Canada.

To date, partnerships have been established through Memorandums of Understanding to provide a national data management plan web service and to develop a preservation pipeline:

- *DMP Assistant*: A project to implement a national bilingual data management planning tool served as the first instance of the envisioned Portage network of expertise. It has been a model of collaboration drawing on a host institution (the University of Alberta), individuals at institutions across the country (the Data Management Planning Expert Group described above), and internationally the work of the DCC. Following initial implementation, the University of Alberta and CARL/Portage entered into a Memorandum of Understanding to set expectations for the ongoing operation of the service. DMP Assistant, launched in October 2015, allows the creation of national templates to meet specific requirements of funding bodies or customised templates for individual institutional use. The bilingual, online web service is available to all researchers in Canada. It is hosted by the University of Alberta Libraries and is based on an implementation of the DCC DMP Online tool. One of the Canadian research funding agencies is working with Portage on a demonstration project involving an identified set of researchers using DMP Assistant, to be carried out in 2016 in conjunction with the adoption of the *Statement of Principles on Digital Data Management*. The Data Management Planning Expert Group is currently developing a help desk ticketing service to accompany the web service, and working with the DCC and the California Digital Library (CDL) to create a unified codebase for DMP Online and CDL's DMPTool.
- *Preservation Pipeline*: In the infrastructure pilot preceding the Portage Network, an experimental preservation pipeline was assembled from Archivematica, Globus Publication, and customised code to establish an integrated workflow. Processing datasets from Canadian-funded research in the recent International Polar Year, this software stack demonstrated that automated processes could generate archival digital objects for research datasets and that these objects could be deposited with an access platform and archived on preservation storage. Once ingested into a discovery and access platform, datasets were discoverable and retrievable under appropriate controlled access conditions. This test also identified several enhancements required for a production system based on this specific design. Coming out of this project was an agreement to build a production service based on improvements to this project's model, formalized in a Memorandum of Understanding between Compute Canada and CARL/Portage in January 2016. Work has begun on software specifications. Discussions have been started around partnerships with other organizations as well, to support data discovery services, preservation storage, enhancements to Archivematica, and middleware between Dataverse and the preservation pipeline.

Stakeholder engagement

Research data management is a collaborative endeavour. The Portage Network involves multiple stakeholders and a primary objective is to build a community of practice for research data management. Developing the two major components of the network indicated above is based on a strong understanding of researchers' needs and solid relationships with funding agencies, data stewards, infrastructure providers, academic library consortia and international collaborators. The first priorities of the Portage Director included developing closer relationships with numerous stakeholder groups, many of whom are represented on the Portage Steering Committee.

Governance and Operational Guidelines

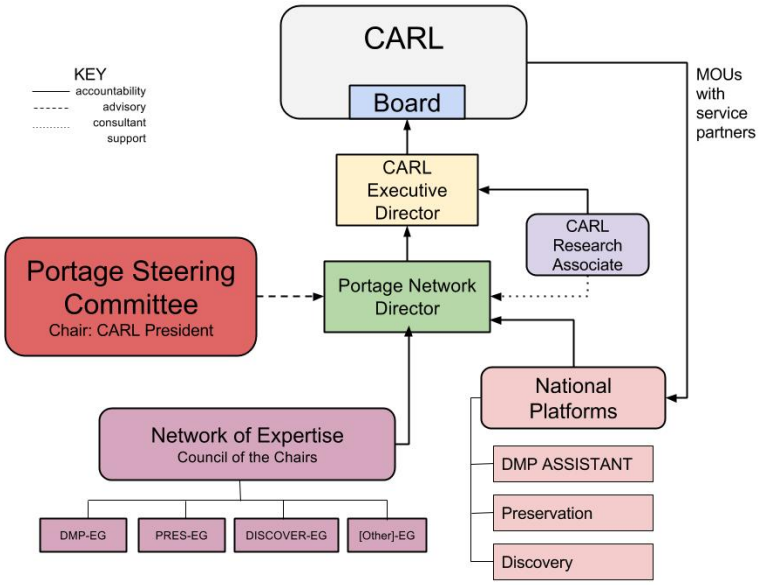

Figure 3: Organizational Chart for the Portage Network, 2016-2017

The Portage Steering Committee will advise on directions and priorities during the two-year development period when operational services are established for the network of expertise and collaborative platforms are developed for planning, preserving and discovering research data. The Steering Committee will also guide plans for Portage beyond this development period, specifically helping the Portage Director prepare both governance and business models for ongoing operations. As well, it will review operating principles, policies, and procedures and will identify priorities for investment and areas for development. The Steering Committee members are representatives of CARL, the regional library consortia, organizations with whom Portage has a formal agreement, national organizations representing research data management services and professionals from universities (including research administrators, research ethics boards, VPRs, and IT services), RDC, the LCDI, and the DCC.

The current operational guidelines for the network have been defined as follows:

1. CARL and its partners have governance roles in the Portage network;
2. Portage, as one component in a larger digital research infrastructure that supports research data management in Canada, will function within this larger context and collaborate with other stakeholders to develop a sustainable and coherent national research data management environment;
3. Significant in-kind contributions will be made by university libraries. As an extension of their operations, participating libraries will provide infrastructure and staff support for both the network of expertise and national platforms for data management plans, preservation, and discovery;
4. Portage will work with other research digital infrastructure providers, such as Compute Canada, CANARIE, and other institutions, to establish in-kind storage capacity for research data and to provide support for national platforms for data management plans, preservation, and discovery;
5. Administrative support for Portage and the Portage Director will be coordinated through the CARL office;

6. Portage will coordinate with other institutions and research organizations the connection of their data repositories to Portage's national platforms through the adherence to community-based standards.

Looking ahead

As Portage moves from project to operational network, a key to its success will be to continue to encourage the ongoing engagement of a broad range of stakeholders. All stakeholders need to see that the network advances their own interests. This includes the regional library consortia, and also the other digital research infrastructure stakeholders represented in the LCDI, all of whom are crucial to the successful delivery of services. Library stakeholders have expressed a desire for a network governance model that remains lightweight and recognizes that the library landscape in Canada is already replete with membership organizations. At the broader level, in seeking funding to sustain its operations for the longer term Portage will need to coordinate its efforts with other stakeholders in the digital research infrastructure landscape. Indeed, Portage cannot be successful in advancing the vision of a distributed infrastructure model without others being successful as well, and without coordination at the level of the LCDI.

Collaboration in Canada usually begins with an eye to what has already been accomplished or is being developed elsewhere, and Portage will continue to build upon and coordinate with work done not only in Canada but also internationally by groups such as the DCC, the Research Data Alliance (RDA) with the development of global data repository certification criteria, and the Confederation of Open Access Repositories (COAR) which is beginning to consider the next generation of research repositories [COAR, 2015]. As Herbert Van de Sompel has recently observed, even when interoperability is the goal and good will abounds, there can be false starts and valid, competing considerations [Van de Sompel, 2015].

Conclusion

In Canada, the development of RDM services, and the broader digital research infrastructure landscape in which they sit, has begun to advance in recent years through the bottom-up collaboration of multiple stakeholders. The Portage Network is developing coherent strategies for sound RDM practices, as it draws on expertise across the country and works to connect the various infrastructure and service components needed for planning, preserving and discovering research data. Portage is contributing to the coordination of digital research infrastructure by filling gaps where tools are missing and bridging systems where interoperability is needed. At the same time, it relies extensively on other key players. At the national level, those players are the groups represented on the LCDI: leaders in research services, research libraries, advanced research computing, networks, administrative standards, research granting councils and researchers themselves. At the institutional level, those same groups come together through various offices and we are seeing new collaborations develop, particularly amongst libraries, research services offices and IT units. There is growing recognition of the many touch points in the research data life cycle that involve librarians, and the services being developed through Portage. With ongoing multi-stakeholder engagement, it is hoped that Portage will continue to help enable institutions to build their capacity for RDM services and contribute to the development of an advanced national digital research infrastructure, and thus increase the impact of the research being conducted across the country and around the world.

Acknowledgements

Numerous people have contributed to the development of the Portage Network. Special thanks to Alan Darnell, Mark Leggott, Steve Marks, Dugan O'Neil, Brian Owen and Leanne Trimble for the detailed development of many of the network concepts, the expert groups for their work on the development of services, CARL's Research Associate Kathleen Shearer for her role in managing the pilot project leading to Portage and her writing on many aspects of the network, and CARL staff member Lise Brin for her work on the Portage website and documentation. Most of all, thank you to Portage Director Chuck Humphrey for his leadership and support from the early days of initial ideas through to the current period of implementation.

Biographical Notes

Martha Whitehead is Vice-Provost and University Librarian at Queen's University, Canada. She is currently President of the Canadian Association of Research Libraries and led the working group that developed the vision and service concepts for the Portage Network. She serves on several national multi-stakeholder bodies, including Research Data Canada's Steering Committee, the Leadership Council for Digital Infrastructure, and the Social Sciences and Humanities Research Council's Programs and Quality Committee.

Donna Bourne-Tyson is the University Librarian at Dalhousie University. Research interests include digital scholarship, OA publishing, and the impact of technology on equitable access, reading and learning. Donna is the Vice-President / President-Elect for the Canadian Association of Research Libraries (CARL), and a founding member of Portage, the national RDM network supported by CARL. She has served as Board Vice-Chair/Chair of the Executive for the Canadian Research Knowledge Network (CRKN), and represented CRKN on the Leadership Council for Digital Infrastructure.

References

Canadian Association of Research Libraries/Association des bibliotheques du recherche du Canada (2013). Retrieved from http://www.carl-abrc.ca/doc/CARL%20Big%20Data%20Consultation%20Response%20Dec%2012%202013.pdf

Confederation of Open Access Repositories (COAR) (2015). Promoting open access and open science: report of the current state of repositories. Retrieved from https://www.coar-repositories.org/files/COAR-State-of-Repositories-May-2015-final.pdf

Davidson, J. (2014) "Emerging good practice in managing research data and research information within UK Universities," Research Data Canada, March 4, 2014. http://www.rdc-drc.ca/activities/webinars/

Humphrey, C., Shearer, K. & Whitehead, M. (2016). Towards a Collaborative National Research Data Management Network. Paper presented at: 11th International Digital Curation Conference.

Humphrey, C. (2016). Portage Status Report. Retrieved from https://portagenetwork.ca/wp-content/uploads/2016/04/Portage_Status_Report_2016_04_EN.pdf

Leadership Council for Digital Infrastructure (2014). Summary report – Digital Infrastructure Summit 2014. Retrieved from http://digitalleadership.ca/wp-content/uploads/2014/02/Summary-Report-of-Summit-2014-Final-March-2014.pdf.

Leadership Council for Digital Infrastructure. (2015). Response to Industry Canada's 2015 Consultation: Developing a Digital Research Infrastructure Strategy.

Open Government Canada. (2014). *Canada's Action Plan on Open Government 2014-16* (Catalogue No. B22-130/2014E-PDF). Retrieved from Open Government Portal: http://open.canada.ca/en/content/canadas-action-plan-open-government-2014-15

Research Data Canada (2016). Research Data Management in Canadian Universities: A Statement of Principles. Retrieved from www.rdc-drc.ca/?wpdmdl=709

Sallans, A. (2016). Strategies for improving openness and reproducibility. Retrieved from http://www.dcc.ac.uk/sites/default/files/documents/IDCC16/Keynotes/Andrew%20Sallans%20Day%201.pdf

TCPS2 (2014). Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. Retrieved from http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/Default/

Tri-Agency Open Access Policy on Publications (2015). Retrieved from http://www.science.gc.ca/default.asp?lang=En&n=F6765465-1

Tri-Agency Statement of Principles on Digital Data Management (2015). Retrieved from http://www.science.gc.ca/default.asp?lang=En&n=83F7624E-1

Universities Canada, 2015. Retrieved from: http://www.univcan.ca/policy-issues/research-and-innovation/

Van de Sompel, H., Nelson, M. Reminiscing about 15 years of interoperability efforts. D-Lib Magazine 21(11/12). Retrieved from http://www.dlib.org/dlib/november15/vandesompel/11vandesompel.html