



Canadian Association
of Research Libraries
Association des bibliothèques
de recherche du Canada

Facilitation, Collaboration and Cooperation: A Canadian Research Data Management Network

A planning meeting hosted by the Canadian Association of Research Libraries
Monday, December 2, 2013
ARC Hotel, Ottawa, ON

Summary

The following notes aim to document the main points of discussion about the proposed national network of library-based research data management services and are not a comprehensive account of the meeting discussion.

Canadian Research Data Management: Gaps, Lessons, Vision

The evolving story of collaborative research data management infrastructure in Canada was presented by Chuck Humphrey (slides available). He outlined three main themes: a strategic shift has occurred over the past decade from building a national data preservation institution to building national research data management infrastructure; building this national research data infrastructure is taking place from the bottom up; building this infrastructure from the bottom up requires intentional, collaborative actions – the driving principle is one of cooperation, not control. Speaking on behalf of Bob McNutt of the Canadian Research Data Centres Network (CRDCN), Chuck also outlined proposed new directions that would include RDCs handling additional kinds of high-confidentiality data. Martha Whitehead, as Chair of the CARL Data Management Subcommittee, briefly reviewed the context of the multiple organizations contributing to developments in the research data management arena, the TC3+ consultation document *Capitalizing on Big Data: Toward a Policy Framework for Advancing Digital Scholarship in Canada*, and the long tail of research data described in the short primer provided by Kathleen Shearer.

Common Ground: Academic Libraries in the Canadian Research Data Management Ecosystem

Individuals from each of the regions (see attached participant list) shared information on initiatives being undertaken consorcially, amongst smaller groups of institutions or at individual institutions. There is experience in each region with digital asset repositories of some form, and all are interested in systems and services to preserve and access research data in response to campus needs and in collaboration with other groups. It was noted that there has been a successful

training and support network amongst data librarians for decades, regionally and nationally, and this model could be broadened. Research Data Canada representatives described the organization's role as an enabler of services and infrastructure across all sectors of the research landscape. CRKN representatives described the Integrated Digital Scholarship Ecosystem concept arising from their recent strategic planning, and its compatibility with the notion of a national research data management network.

Defining the nature of the network

Why Should the Network Exist?

- Research integrity is important, and it requires sound research data management
- Managing research data is too big and complex and too critical to deal with only at the local or regional level
- To avoid silos, it is important for researchers to have linkages across institutions and regions
- To ensure that all researchers, regardless of institution, have access to services and resources – to provide as level a playing field as possible across the country
- If we don't do anything, the data will go into international repositories, to vendors, or stay in researchers' hard drives
- To support anticipated requirements of funding councils and from journals for data management and sharing
- To develop national collections of valuable research data produced in Canada
- To ensure stable, long-term preservation of data: a network provides alternatives that a stand-alone preservation service will find challenging; a network can have a succession plan for the data in case one institution is no longer able to archive it
- A network can provide shared services, expertise and technology
- Why does it not exist now? Because it has taken some time for stakeholders to articulate the meaning of national infrastructure for research data management (the TC3+ consultation on big data engages dialogue around the new paradigm of the pursuit of national infrastructure for research data instead of a national institution); because national funding is geared to a 3-5 year research cycle

What Will the Network Look Like?

- National and inclusive of all academic libraries (beyond CARL)
- The network could be organized like a consortium: either regionally or functionally

- The structure of the network could be built on the regional consortia (CANARIE and Compute Canada are both models of regional groups coming together)
- CRKN could play an administrative role, although at the same time don't want to distract CRKN from their major mission
- Focus on library strengths, sharing services and building on those strengths

What Will The Network Do?

1. Education & Expertise

- Education and training is required in several stakeholder communities: researchers, research support, and librarians
- Provide a locus for expertise that could be called upon by universities to help set up services or assess/auditing existing services (even for a fee)
- The DCC is a possible model: it offers training, but also goes into institutions with teams of people to help with local development of services
- In a large country like Canada, this type of training would need to be very decentralized so that expertise can be close to the researchers
- Must coordinate with information schools in their work on data management in the MLIS curricula
- Data management could be considered a part of information literacy and the network could develop material for education of undergraduates as well as graduate students
- The network could also facilitate research in digital preservation issues

2. Services

- The network could support services in all three of the following areas:
 1. I need access to data that exists; how do I find it and use it (e.g. text mining, visualization, etc)
 2. I am starting new research and need help with planning data management
 3. I didn't have a data management plan and now I'm looking for solutions for the care of data from my finished research project
- A major benefit of the network would be informed referral to expert advice
- The network could provide services in the area of data visualization
- We'll need to ensure that methods around metadata and related support will scale
- Data rescue might also be a role for the network: a team could be deployed on data rescue missions – for when projects and programs end or get shutdown quickly
- The network could also play a role in data modeling

3. Tools and Technology

- Offer a preservation platform, perhaps working with Compute Canada for storage and also offering other specialized storage that might be required

- The libraries' expertise is in the management of the content
- Focus on dissemination, discovery, preservation, and access
- Specialized storage
- New "stacks" for the future for both researcher-produced and licensed data
- Commodity equipment and open source
- Technology could be offered through a service provider model of regional nodes
- Identify and disseminate standards for preservation storage
- Establish requirements for data integrity (need to ensure that data is readable/analyzable by software and machines into the future).
- Manage a process for certifying data repositories

Challenges

- Scale and scope: need to define what can be done now, how it could grow or intersect with other initiatives and what is needed to make that happen
- Long-term policy: are the participating universities and their libraries taking responsibility for these data forever
- Issues are complex and require coordination across a number of stakeholders communities
- Coordination is needed across provinces and regions
- Libraries have made progress but need to further engage researchers, VP's Research and other administrators
- There may be a backlash from researchers when they realize that data management activities could cut into actual research funding
- Currently, there are no obvious funding programs (although 3 billion per year spent on research each year in Canada – could some be targeted towards research data management)
- Collaboration requires that participants commit resources: intentional interdependency

Core Values

- Respect for privacy and intellectual property
- Collaborative approaches: cost savings and shared expertise
- Serving those without the ability to pay
- Open access: data as a public good
- Intelligent access: openness—with respect for privacy
- Interoperability
- International connections
- Stewardship: sense of responsibility for valuable assets
- Not one-size fits all (data from different disciplines may have to be treated differently)

Next Steps

Based on work already under way, a national coordinating network on even a relatively small scale would be useful now. It should be developed in such a way that it can scale up, and that it can integrate with and potentially provide a framework for larger-scale research data management services in the future.

- Form a representative group to develop a project charter drawing on the day's discussions
- Identify a project coordinator (CARL has allocated some funding)
- The charter and governance documents will come back to potential network participants, either through regional representation or individually

Facilitation, Collaboration and Cooperation: A Canadian Research Data Management Network

A planning meeting hosted by the Canadian Association of Research Libraries – December 2, 2013

Participants

Clare Appavoo	Executive Director, Canadian Research Knowledge Network (CRKN)
Allan Bell	Director, Library Digital Initiatives, University of British Columbia
Jonathan Bengtson	University Librarian, University of Victoria; incoming Chair, COPPUL
Pam Bjornson	Director General, Knowledge Management, National Research Council; Chair, Research Data Canada
Donna Bourne-Tyson	University Librarian, Dalhousie University; representative, CAUL; Chair, Executive Committee, CRKN
Amy Buckland	eScholarship, ePublishing & Digitization Coordinator, McGill University Library
Pascal Calarco	Associate University Librarian, Digital & Discovery Services, University of Waterloo
Alan Darnell	Director, Scholars Portal Services, Ontario Council of University Libraries
Michelle Edwards	Data Librarian, University of Guelph (participating remotely)
Margaret Haines	University Librarian, Carleton University; Chair, Ontario Council of University Libraries
Chuck Humphrey	Coordinator, Data Library, University of Alberta
Karen Keiller	Director of Information Services & Systems, University of New Brunswick Libraries; Chair, Council of Atlantic University Libraries
Stephen Marks	Digital Preservation Librarian, Scholars Portal
David Moorman	Senior Programs Planning Officer, Canada Foundation for Innovation
Brian Owen	Associate University Librarian, Processing and Systems, Simon Fraser University
Jennifer Riley	Associate Dean, Digital Initiatives, McGill University Library
Brent Roe	Executive Director, Canadian Association of Research Libraries
Diane Sauvé	Directrice, Soutien à la réussite, à la recherche et à l'enseignement, Direction générale de la Direction des bibliothèques, Université de Montréal
Kathy Scardellato	Executive Director, Ontario Council of University Libraries
Kathleen Shearer	Research Associate, Canadian Association of Research Libraries
Walter Stewart	Coordinator, Research Data Canada
Martha Whitehead	University Librarian, Queen's University; Chair, CARL Data Management Subcommittee; Meeting Facilitator