

# DATA MANAGEMENT ACTIVITIES OF CANADA'S NATIONAL SCIENCE LIBRARY – 2010 UPDATE AND PROSPECTIVE

*Mary Zborowski*

*Strategy and Development Branch (NRC-SDB), National Research Council Canada, Ottawa ON Canada  
K1A 0R6*

*Email: [mary.zborowski@nrc-cnrc.gc.ca](mailto:mary.zborowski@nrc-cnrc.gc.ca)*

## ABSTRACT

*NRC-CISTI serves Canada as its National Science Library (as mandated by Canada's Parliament in 1924) and also provides direct support to researchers of the National Research Council of Canada (NRC). By reason of its mandate, vision, and strategic positioning, NRC-CISTI has been rapidly and effectively mobilizing Canadian stakeholders and resources to become a lead player on both the Canadian national and international scenes in matters relating to the organization and management of scientific research data. In a previous communication (CODATA International Conference, 2008), the orientation of NRC-CISTI towards this objective and its short- and medium-term plans and strategies were presented. Since then, significant milestones have been achieved. This paper presents NRC-CISTI's most recent activities in these areas, which are progressing well alongside a strategic organizational redesign process that is realigning NRC-CISTI's structure, mission, and mandate to better serve its clients. Throughout this transformational phase, activities relating to data management remain vibrant.*

**Keywords:** NRC-CISTI, National Research Council Canada, Scientific and technical research data, Research data management, Canada, National activities

## 1 INTRODUCTION

For over 80 years, the National Research Council's Canada Institute for Scientific and Technical Information (NRC-CISTI) (<http://cisti-icist.nrc-cnrc.gc.ca/eng/ibp/cisti/index.html>) has been an integral component in the growth of Canada's innovation capacity. As Canada's national science library, NRC-CISTI provides Canada's research community with high-value authoritative information and information services in science, technology, and health to accelerate innovation and commercialization.

In this author's previous DSJ paper ([http://www.jstage.jst.go.jp/article/dsj/8/0/8\\_27/article](http://www.jstage.jst.go.jp/article/dsj/8/0/8_27/article)), the nature and positioning of NRC-CISTI within the Canadian national and international scene as well as some of its research data activities and initiatives were discussed. This paper will present some of the successes that have resulted as well as NRC-CISTI's future plans.

## 2 THE NRC-CISTI TRANSFORMATION

On February 18<sup>th</sup>, 2009, NRC announced the results of a strategic review of NRC programs conducted by the Government of Canada. As a result of this review, NRC-CISTI realigned some of its resources and programs. These transformational activities, documented in several media releases and articles,

<http://cisti-icist.nrc-cnrc.gc.ca/eng/ibp/cisti/media-room/index.html>;

<http://cisti-icist.nrc-cnrc.gc.ca/eng/ibp/cisti/newsletters/cisti-news/2009june.html>;

<http://cisti-icist.nrc-cnrc.gc.ca/eng/ibp/cisti/newsletters/cisti-news/2009fall.html>;

<http://cisti-icist.nrc-cnrc.gc.ca/eng/news/cisti/2009/transformation-update.html>;

<http://cisti-icist.nrc-cnrc.gc.ca/eng/news/cisti/2010/nrc-cisti-update.html>, include:

- The journals and services of NRC Research Press were transferred to the private sector (<http://cisti-icist.nrc-cnrc.gc.ca/eng/news/cisti/2010/transformation-research-press.html>). A new not-for-profit corporation was created for the NRC Research Press journals and services: Canadian Science Publishing (<http://pubs.nrc-cnrc.gc.ca/eng/home.html>).
- Strategic partnership initiatives such as PubMed Central Canada, the Research Data Strategy Working Group, the Trusted Digital Repository, and health sector initiatives have been pursued with new purpose and commitment.

- Mechanisms to offer licensed access to information content and in-depth information services to NRC and information discovery and access services to Canadians and researchers from around the world have been refined and streamlined.

More recently, in early 2010, the Government of Canada announced the appointment of Mr. John R. McDougall (<http://www.nrc-cnrc.gc.ca/eng/news/nrc/2010/04/01/new-president-nrc.html>) as the new President of NRC. Mr. McDougall's previous experiences with industry and other R&D supporting agencies have had a major impact on the evolution of public policy in business, trade, science, and technology as well as innovation in both Canada and beyond. His objectives for NRC, which will impact NRC-CISTI strategic directions and activities, include helping to define national research issues and needs.

### **3 UPDATE ON NRC-CISTI'S DATA-RELATED ACTIVITIES**

#### **3.1 The Canadian National Committee for CODATA (CNC/CODATA)**

NRC-CISTI has been the sponsor of this national CODATA committee since its inception and continues to actively support its activities. CNC/CODATA's Chair is an active participant in the Research Data Strategy Working Group (discussed below). CNC/CODATA's activities, such as the Sangster Award for Young Scientists ([http://www.codata.org/canada/sangster/sangster\\_e.shtml](http://www.codata.org/canada/sangster/sangster_e.shtml)) and the Report on Data Activities in Canada (<http://www.codata.org/canada/DAC/>), are only possible due to NRC-CISTI's ongoing support.

#### **3.2 Research Data Strategy Working Group (RDS WG)**

NRC-CISTI continues to lead this Working Group, originally set up to address challenges and issues surrounding the access and preservation of data arising from Canadian research. Chaired by NRC-CISTI's Director General, Pam Bjornson, this multi-disciplinary group of universities, government institutions, research libraries, funding agencies, and research organizations recognizes that there is a pressing need to deal with Canadian data management issues. Its activities focus on the actions and leadership roles that researchers and institutions can take to ensure Canada's research data is accessible and usable for current and future generations of researchers.

A major deliverable of the Working Group ([http://data-donnees.gc.ca/eng/news/gap\\_analysis.html](http://data-donnees.gc.ca/eng/news/gap_analysis.html)) was Stewardship of Research Data in Canada: A Gap Analysis (<http://data-donnees.gc.ca/docs/GapAnalysis.pdf>) (Dec. 2009). Using the data lifecycle as a framework, the report examined Canada's current state versus an 'ideal state' based on existing international best practices across 10 indicators, including policies, funding, roles and responsibilities, standards, data repositories, skills and training, accessibility, and preservation. The analysis revealed significant barriers to the access and preservation of research data — barriers that could have a serious impact on the future of Canadian research and innovation if not addressed. For example, large amounts of data are being lost because of the woefully inadequate number of trusted data repositories in Canada. The results of this gap analysis provide further evidence of the need to address these pressing issues and will feed the development of a national strategy for improving data stewardship in Canada.

The Working Group also supported a major deliverable of the Canadian Association of Research Libraries: the awareness toolkit, Research Data: Unseen Opportunities ([http://www.carl-abrc.ca/about/working\\_groups/pdf/data\\_mgt\\_toolkit.pdf](http://www.carl-abrc.ca/about/working_groups/pdf/data_mgt_toolkit.pdf)). This toolkit provides background information about the management of research data in Canada and recommended directions to be taken. A companion document will address roles for libraries in this area to provide readers with a general understanding of the current state of research data in Canada and internationally.

This year, the Government of Canada unveiled a national Digital Economy Consultation (<http://de-en.gc.ca/>), aimed at building consensus among governments, the private sector, academia, and the Canadian public in developing a digital economy strategy for Canada. The commitment is aimed at positioning Canada for leadership in the global digital economy. The key themes being considered include: Capacity to Innovate Using Digital Technologies; Building a World-Class Digital Infrastructure; Growing the Information and Communications Technology Industry; Digital Media: Creating Canada's Digital Content Advantage; and Building Digital Skills for Tomorrow.

The RDS WG's formal response to this consultation <http://data-donnees.cisti-icist.nrc-cnrc.gc.ca/gsi/ctrl?lang=e> highlighted the fundamental importance of ensuring that Canada's research data is accessible over the long term and stressed the need for a coordinated national strategy for Canada, especially in view of advances in this area

being made by other countries. It proposed five recommendations, outlining what Canada can do to ensure full economic benefits are being derived from the research data created by governments and through publicly funded research:

- Implement a national approach to research data;
- Strengthen Canada's existing infrastructure of data centers, data repositories, and data libraries;
- Encourage data stewardship policies in granting councils and foundations, government departments, and universities and colleges;
- Form a national oversight mechanism that would provide support for data stewardship activities in Canada;
- Launch an open data initiative in Canada.

The Working Group has been collaborating with CNC/CODATA to develop data management courses. While CNC/CODATA is focusing on a course for researchers, the Working Group's course is aimed at developing data-management skills for librarians. A pilot version of the course was fully attended and held in November 2009. A resource pack ([http://www.library.carleton.ca/about/projects/data\\_seminar/resources.html](http://www.library.carleton.ca/about/projects/data_seminar/resources.html)) was prepared for all participants and is available on the web. The seminar was repeated in spring of 2010 as part of the Data Liberation Initiative (<http://www.statcan.gc.ca/dli-ild/dli-idd-eng.htm>). The success of this Initiative has made it clear that Canadian data professionals are anxious to learn more about research data management and are willing to start working with researchers in this area. While the introductory seminar is essential in providing participants with grounding on the issues, the evaluations have indicated that most participants eagerly await the companion "best practices" module in preparation by CNC/CODATA.

The Working Group is additionally working on a set of Canadian Data Stewardship Principles that will expand on the 2004 OECD Ministerial Declaration on Access to Publicly Funded Research Data. This work is expected to be finished in late 2010 and is likely to include the following pillar concepts:

- Research data are deemed valuable global assets. Data need to be shared to support the transparency of discovery, allow replication of published findings, and facilitate creation of new data from existing data.
- Innovations in IT are vital to managing, providing access to, and preserving research data. Research data and corresponding metadata need to be produced using standards designed for interoperability and prepared using quality control and secure practices.
- Stakeholder collaboration is required for successful data stewardship. It takes a research community to manage and preserve its data. Stakeholder roles and responsibilities must be clearly articulated and institutional mandates developed.
- Data stewardship is built on shared values in communities of practice. The skills and norms of good data management and stewardship require the involvement of communities of practice within domains of research.

### 3.3 NRC-CISTI's Gateway to Scientific Data

Scientific data generated during the research process can be an important resource for researchers, but only if it is accessible and usable. In 2009, NRC-CISTI, working in collaboration with CNC/CODATA and drawing from its Report on Data Activities in Canada, launched ( <http://cisti-icist.nrc-cnrc.gc.ca/eng/news/cisti/2009/gateway-launch.html> ) the Gateway to Scientific Data (<http://data-donnees.cisti-icist.nrc-cnrc.gc.ca/gsi/ctrl?lang=en> ) whereby researchers now have a central gateway for easier links to Canadian scientific, technical, and medical (STM) data sets and other important data repositories. The Gateway will help ensure that the valuable data generated by Canadian researchers is more easily accessible for reuse in other research endeavours. With the ability to access and use data from a multitude of sources, researchers will be better positioned to turn research into discoveries and innovations. Besides offering access to such data resources, the Gateway also provides links to selected best practices and policies guiding data management and curation activities in Canada and links to selected journals and upcoming conferences and meetings.

This new initiative is part of NRC-CISTI's contribution to a broader national initiative undertaken by the RDS Working Group to address the challenges and issues surrounding the access and preservation of data arising from Canadian research and will provide researchers with access to the information resources they need to remain at the forefront of R&D developments in Canada and around the world.

### 3.4 ICSTI conference 2009: Managing Data for Science

In June 2009, NRC-CISTI hosted the annual conference of the International Council for Scientific and Technical Information (ICSTI), which carried the theme "Managing Data for Science". Fourteen speakers and

over 140 participants discussed issues relating to how researchers, librarians, and publishers can work together to create structures to manage and communicate scientific data. Highlights included notable presentations on particularly relevant topics such as:

- Where the responsibility lies concerning the preservation of research data – with the researcher or with the librarian;
- New approaches to encourage scientists to share their data, in particular, assigning Digital Object Identifiers (DOIs) to datasets, in turn allowing the data to be cited as if it were a regular publication;
- Project concepts to enable access and online analysis of data from various sources.

### 3.5 DataCite

DataCite (<http://www.datacite.org/>) is an international consortium of 12 libraries and research centres from nine countries committed to increasing access to research data on the internet and ensuring a better return on investment of publicly funded research. DataCite is a global registration agency for research data whose members work with datacentres to assign digital object identifiers (DOIs) to their datasets. DOIs provide researchers with a method to locate, identify, and cite research datasets over the long term and serve to increase the acceptance of research data as legitimate, citable contributions to the scientific record.

As a full member of DataCite, NRC-CISTI will soon develop the necessary technical infrastructure to establish a data registration centre for Canadian datasets. NRC-CISTI has also received funding from the Government of Canada's Program for International Polar Year (IPY) to register and assign DOIs to Canadian data collected during the 2007-2008 IPY, the largest-ever international program of coordinated, interdisciplinary science focused on Polar Regions. The benefits for researchers will include more efficient research, fewer resources spent duplicating experiments, greater collaboration across disciplines and sectors, and greater researcher recognition via data citations.

## 4 ADDITIONAL CONTEXT

NRC-CISTI is involved in several other initiatives which, while not directly related to the management of research data, do effectively illustrate NRC-CISTI's positioning within the Canadian national and international contexts.

### 4.1 Federal Science eLibrary (FSeL)

For several years, NRC-CISTI has been working with five science-based Canadian federal departments and agencies (collectively known as the (Canadian) Strategic Alliance of Federal Science and Technology Libraries) to explore improvements and efficiencies in the delivery of full-text electronic journals in science, technology, and medicine (STM) to the desktops of all federal government researchers, policy analysts, and decision makers. The initiative has sought new government funding to transform the current process used for acquiring access rights to electronic journals for all federal departments, thereby ensuring all employees have desktop access to the information they need to work efficiently and productively.

Success has now been realized with the signing of the first agreement with the key scientific publisher Springer (<http://safstl-asbstf.scitech.gc.ca/eng/news/2010-04-29-fsel-licence-agreement.html>). By this agreement, access to essential STM research has been expanded for significantly less money than before, by effective leveraging of existing library negotiation expertise with the S&T publishing sector. The next steps include securing similar agreements with other publishers to expand access even further. This agreement is an important step forward in making the FSeL a reality. It demonstrates the power of collective negotiation by federal libraries and will serve as a test case for broader implementation for other science-based departments and agencies (SBDAs). The vision for the FSeL remains the same: to ensure seamless and equitable desktop access to STM e-journal for all federal employees.

### 4.2 WorldWideScience.org

NRC-CISTI is a founding member of the WorldWideScience Alliance, which has developed a new multilingual translation tool to provide the first-ever, real-time searching and translation of globally dispersed multilingual scientific literature. The launch (<http://cisti-icist.nrc-cnrc.gc.ca/eng/news/cisti/2009/worldwidescience.html>; <http://cisti-icist.nrc-cnrc.gc.ca/eng/news/cisti/2010/multilingual-translation-tool.html>) of this tool broadens access to global science, allowing users to search databases in China, Russia, France, and several Latin

American countries and receive the search results translated into one of nine languages (Chinese, English, French, German, Japanese, Korean, Portuguese, Spanish, and Russian – with more languages to be added in coming months). It uses a federated search engine technology to search scientific and technical research and development data, including development data from close to 80 percent of the world's population-information which, prior to its availability from WWS, was previously not widely accessible through common search engines. WWS has experienced exponential growth – from 10 countries at the start to 56 as of 2010 – and currently places more than 375 million pages of scientific information at the world's fingertips.

NRC-CISTI's Director General, Pam Bjornson, is Vice Chair of WWS and one of its original charter signatories.

### **4.3 Canadian Health Sciences Community**

NRC-CISTI has been an integral part of the Health Sciences community since 1966 when, in addition to being Canada's National Science Library, it was also recognized as the National Library of Health Sciences for Canada. NRC-CISTI leverages services and infrastructure already developed for NRC itself. Two prime examples of these sorts of initiatives are PubMed Central Canada (PMC Canada) and the Canadian Virtual Health Library.

#### **4.3.1 PubMed Central Canada (PMC Canada)**

PMC Canada is the result of a partnership between NRC-CISTI, the Canadian Institutes of Health Research (CIHR), and the US National Library of Medicine (NLM). With its launch (<http://cisti-icist.nrc-cnrc.gc.ca/eng/news/cisti/2010/pubmed-central-canada-launch.html>), Canadians have free access to a national digital repository of the latest peer-reviewed health and life sciences literature, including research resulting from CIHR funding. Additionally, Canadian researchers will now contribute to a growing, searchable digital archive of published Canadian health research.

PMC Canada supports CIHR's Policy on Access to Research Outputs, which requires CIHR grant recipients to make their peer-reviewed publications freely accessible online within six months of publication. PMC Canada's manuscript submission system will enable CIHR-funded researchers to deposit their peer-reviewed articles, exposing their research to a global audience and facilitating collaboration to advance scientific progress.

This first phase of PMC Canada includes a bilingual interface, a manuscript submission system for CIHR researchers, and a bilingual help desk. Plans for the second phase of the repository will incorporate a customized web front-end along with enhanced reporting and alerting features for system funders and users. An advisory committee of Canadian health researchers and other stakeholders will guide PMC Canada's future development.

PMC Canada builds on PubMed Central, the archive developed by the US National Library of Medicine and joins UK PubMed Central (UKPMC) as a member of the larger PMC International network. PMC Canada will connect researchers to this international network, making much of PMC and UKPMC content accessible to them. Such international cooperation will improve worldwide access to life sciences literature and scientific data, aiding the progress of science and improving public access to the results of research that affects health and health care.

#### **4.3.2 Canadian Virtual Health Library (CVHL)**

In July 2010, CIHR and the Canadian Health Libraries Association (CHLA) launched (<http://cisti-icist.nrc-cnrc.gc.ca/eng/news/cisti/2010/canadian-virtual-health-library.html>) the Canadian Virtual Health Library (<http://chla-absc.ca/nnlh/cvhl/>), which will improve Canadian healthcare by making high-quality health information and evidence easily accessible to all health care professionals, wherever they are in Canada, so that they can make informed decisions about patient care and education and policy and program development.

Rolled out over a three-year period, the CVHL will build a network to link Canadian local, regional, and provincial health library services and maximize resources through coordinated resource sharing and licensing. A bilingual web portal will provide an easy entry point to these resources and to library expertise, resulting in a tailored, rapid response to health professionals' requirements for evidence-based information to support patient care decisions, health promotion, and public health programming. NRC-CISTI has been deeply involved with the CVHL initiative since its inception, a reflection of the Canadian federal government's priority with regards to getting the latest evidence-based information and resources into the hands of those responsible for patient care. Such coordinated access to information resources and services will promote utilization of best practices

and support all health care professionals in fulfilling their mandate to enhance patient outcomes and provide safe, competent, ethical care.

#### 4.4 ICSTI

NRC-CISTI's Director General currently serves as Treasurer of ICSTI (<http://www.icsti.org/>) the International Council for Scientific and Technical Information. As mentioned, NRC-CISTI hosted the 2009 conference, "Managing Data for Science", in Ottawa, Canada. ICSTI offers a unique forum for interaction between organizations that create, disseminate, and use scientific and technical information, and its mission cuts across scientific and technical disciplines as well as international borders to give member organizations the benefit of a truly global community. Maintaining active involvement with ICSTI allows NRC-CISTI to move forward in its research data management activities with a balance appropriate to its role as Canada's National Science Library.

### 5 THE NEXT FEW YEARS

In the next few years, NRC-CISTI will continue development and support of the initiatives described above, by leveraging the services and infrastructure it develops for NRC to expand access to essential research information and data for Canada's research, innovation, and health communities as well as by providing long-term access to the record of Canadian science and federally funded research outputs. Its focus will be directed to the following key areas:

- Working with other federal science departments and the Canadian health community to provide equitable access to scientific, technical, and health information. For example, a new partnership with Health Canada to share library services was established in April 2010. This complements the partnership NRC-CISTI has had for many years with Agriculture and Agri-Food Canada for document delivery services.
- Building access vehicles that showcase Canada's research. In particular, the NRC Publications Archive, a citation and full-text database of NRC-authored research, will continue to be developed, refined, and promoted.
- Expanding access to research data, including ongoing expansion of the Gateway to Scientific Data and support to CNC/CODATA.

In addition, several key initiatives are under way:

#### 5.1 A Data Strategy Framework and Policy for NRC-CISTI

The National Research Council of Canada is itself a producer of scientific knowledge and data, and NRC-CISTI is currently re-examining its role to promote and make accessible NRC's own scientific data sets. NRC-CISTI is currently in the process of developing a long-term data strategy framework, which will address:

- Providing dataset management and identity services;
- Improving access to and discovery of scientific datasets;
- Establishing NRC as a government data provider; and
- Sustaining NRC-CISTI's outreach to the Canadian scientific data research community.

Although NRC-CISTI's activities in these areas are already well-established, an environmental scan and SWOT (Strength, Weakness, Opportunity, and Threat) analysis were undertaken in support of development of a formal policy, expected in late 2011, which will address both NRC needs and those of the national and international stakeholders supported by NRC-CISTI.

#### 5.2 A Data Summit in early 2011

The Research Data Strategy Working Group is proposing to convene a "Research Data and Innovation Summit" in early 2011. The summit will bring together between 30 and 50 senior decision-makers from government, academia, and the NGO and private sectors. Its purpose will be to develop a shared vision and action plan that can provide a framework for future initiatives and assist us all to leverage scarce resources for maximum impact. The anticipated outcomes of the summit include the development of:

- A consensus around some of the key actions required to develop a national strategy;
- A shared vision of the roles and responsibilities of stakeholders in order to ensure that research data management responsibilities are coordinated and effective; and
- New collaborations and working relationships amongst the various participants of the summit.

More information on the Data Summit will be communicated when available on both the NRC-CISTI and RDS WG web sites.

## 6 CONCLUSION

Implementing the transformation required by the Government of Canada Strategic Review is a major focus for NRC-CISTI right now. The scale of this change is huge and very complex and will result in a more focused NRC and national science library. Although the mechanisms and timelines moving forward may still be subject to change, the work on strategic initiatives and collaborations described above will remain integral to the new NRC-CISTI.

## 7 ACKNOWLEDGEMENTS AND RESOURCES

The author thanks her NRC-CISTI colleagues, Alison Cachia, Karen Morgenroth, André Vellino, and Bronwen Woods, for their valuable input as well as John Broome (Natural Resources Canada), Chair of CNC/CODATA. More information may be obtained from the sources below.

CISTI web site: <http://cisti-icist.nrc-cnrc.gc.ca/>

CISTI's Key Strategic Activities: <http://cisti-icist.nrc-cnrc.gc.ca/eng/ibp/cisti/about/overview-initiatives.html>

CISTI's Media Room (links to News, Press Releases): <http://cisti-icist.nrc-cnrc.gc.ca/eng/ibp/cisti/media-room/index.html>

Ms. Pam Bjornson, Director General, NRC-CISTI: <http://cisti-icist.nrc-cnrc.gc.ca/eng/ibp/cisti/biographies/pbjornson.html>

Canadian National Committee for CODATA: <http://www.codata.org/canada>

Documents and links inspiring the work of the RDS Working Group: [http://data-donnees.gc.ca/eng/docs\\_links.html](http://data-donnees.gc.ca/eng/docs_links.html)

RDS Working Group Task Groups and Leads: <http://data-donnees.gc.ca/eng/groups/taskgroups.html>

Zborowski, Mary. *CISTI'S Activities in Support of Scientific Data Management in Canada 2008-2010*. Data Science Journal, Vol. 8 (2009) pp. 27-33. [http://www.jstage.jst.go.jp/article/dsj/8/0/8\\_27/\\_article](http://www.jstage.jst.go.jp/article/dsj/8/0/8_27/_article)

(Article history: Received 15 November 2010, Accepted 6 December 2010, Available online 13 January 2011)