# Research Data : Unseen Opportunities

An Awareness Toolkit commissioned by
the Canadian Association of Research Libraries (CARL)

This toolkit, developed for the **CARL Data Management Sub-Committee**, provides background information about the management of research data in Canada. The purpose is to enable Canadian Association of Research Libraries Directors to raise awareness of the issues of data management with administrators and researchers on campus. The toolkit, along with a companion document that discusses roles for libraries in the management of research data, will provide readers with a general understanding of the current state of research data in Canada and internationally, and offer suggestions to begin addressing the issues involved.

By Kathleen Shearer

**CARL Data Management Sub-Committee**
Marnie Swanson (Chair, University of Victoria)
Pam Bjornson (NRC-CISTI)
Lynn Copeland (Simon Fraser University)
Michelle Edwards (University of Guelph)
Bernie Gloyn (Observer, Statistics Canada)
Margaret Haines (Carleton University)
Janine Schmidt (McGill University)
Kathleen Shearer ( Research Associate, CARL)
Diego Argáez (Secretary, CARL)

Thanks to Diane Sauvé (Université de Montréal) for assisting with the French translation.




CARL ABRC

# Research Data: Unseen Opportunities

# Contents

# The 'Big Picture'

Research data are defined here as the factual records (e.g. microarray, numerical and textual records, images and sounds, etc.) used as primary sources for research, and that are commonly accepted in the research community as necessary to validate research findings.

The US National Science Foundation describes digital data as both the products of research and, increasingly, the starting point for new research. In the digital environment, it is possible for researchers to re-purpose data – to use them in innovative ways and combinations not envisioned by those who created them. It is also common for researchers to use data from one study to build on and enhance the findings of previous research; and to undertake longitudinal studies that compare data from repeated observations of the same items over long periods of time — often many decades [1].

Over the next five years, the world will produce more research data than has been created in all of human history [2].

For the most part these data are born digital, and stored and managed electronically, making them easy to share, replicate, and combine with other data. However, in order to share and reuse data, they must be created and maintained in a manner consistent with the goal of long-term preservation. This involves active data management throughout the life-cycle of the data, beginning at the time they are first envisioned.

Right now in Canada, the vast majority of research data is being lost. For example, a study of Social Sciences and Humanities Research Council funded research projects found that only 3 out of 110 studies had archived their data in a repository, and those 3 were all housed in the US [3]. Research data in Canada are not being systematically managed and therefore, valuable data are under-utilized. While certain disciplines and research projects have institutional, national or international support for data management, this support is available for a minority of researchers only.

A crucial aspect to creating data with long-lasting usability is to ensure that the accompanying documentation is user-friendly, clear, and comprehensive. Ideally, metadata and documentation should be produced at the start of a research project and enhanced throughout the course of the data life-cycle. Planning and support from data professions at the initial stages of the research project can significantly reduce the time and money needed to provide long-term access.

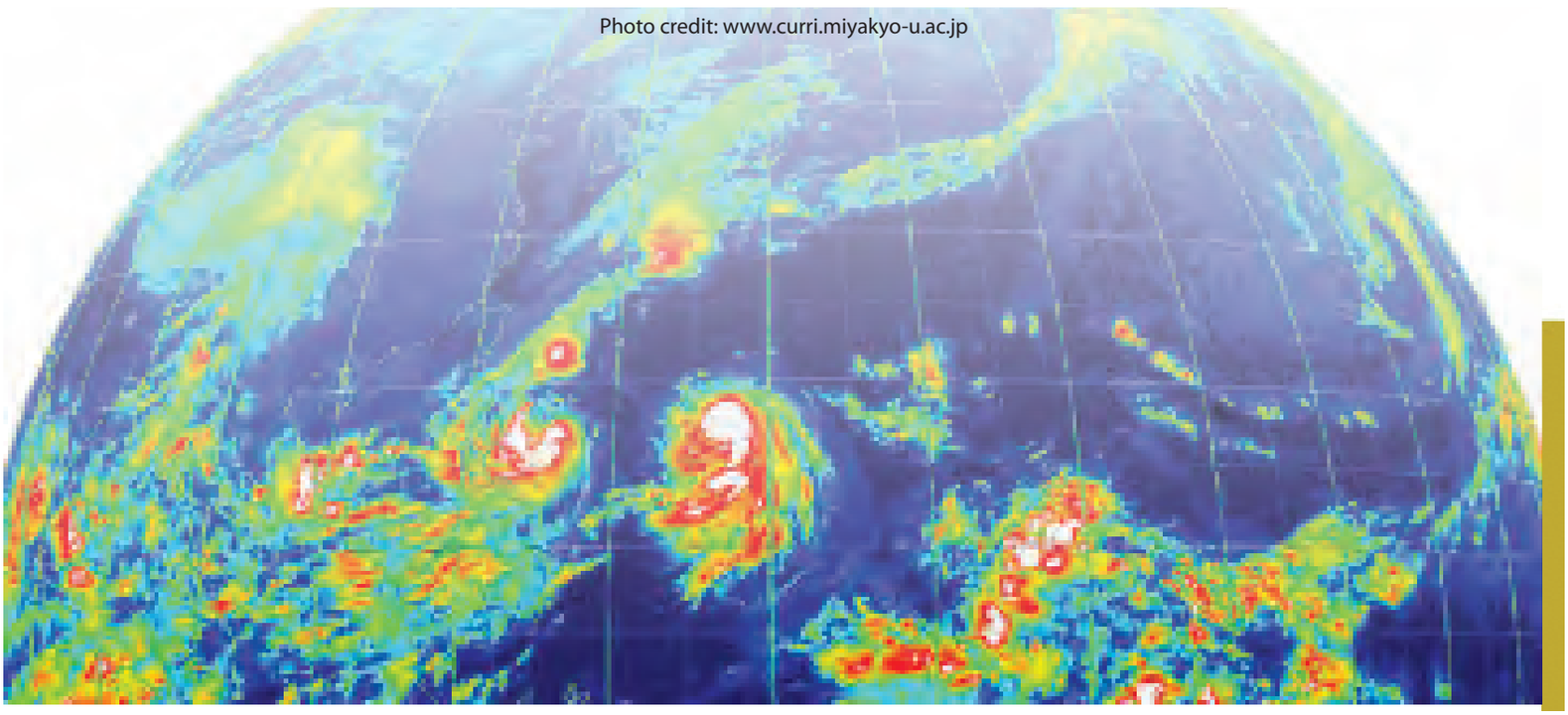
---

[1] "Community-based Data Interoperability Networks". 2009. National Science Foundation. http://www.nsf.gov/pubs/2007/nsf07565/nsf07565.htm

[2] Beagrie, N. "Digital Preservation: Setting the Course for a Decade of Change". 2007. www.abd-bvd.be/documents/60/BDA-1107_NB.ppt

[3] Humphrey, C. "Preserving research data: A time for action". In Preservation of electronic records: new knowledge and decision-making. 2003. Canadian Conservation Institute: Ottawa. pg. 83-89.

Photo credit: www.curri.miyakyo-u.ac.jp



# Major Benefits of Data Management, Sharing, and Reuse

**Accelerates scientific progress.** The sound management of research data will allow researchers to access and understand others' data and re-use them for their own scientific purposes, thereby speeding up the rate of new discoveries.

**Increases the visibility and impact of research.** Data made visible through a data repository can dramatically increase the impact of that research. Sharing research data has been associated with increased citation rates. For example, a study of citation rates for cancer clinical trials publications found that clinical trials that shared their data were cited about 70% more frequently than clinical trials that did not.[4]

**Ensures compliance with funding agency policies.** A growing number of funding agencies demand that researchers and host institutions retain, manage and share their data upon completion of a research project. Thus, universities and researchers have legal and ethical obligations to provide a legacy of research data. Some publishers also require that the data connected to their publications are preserved.

**Avoids duplication of research.** When a dataset is publicly available it is much less likely to be recreated, avoiding expensive and needless data collection/production activities.

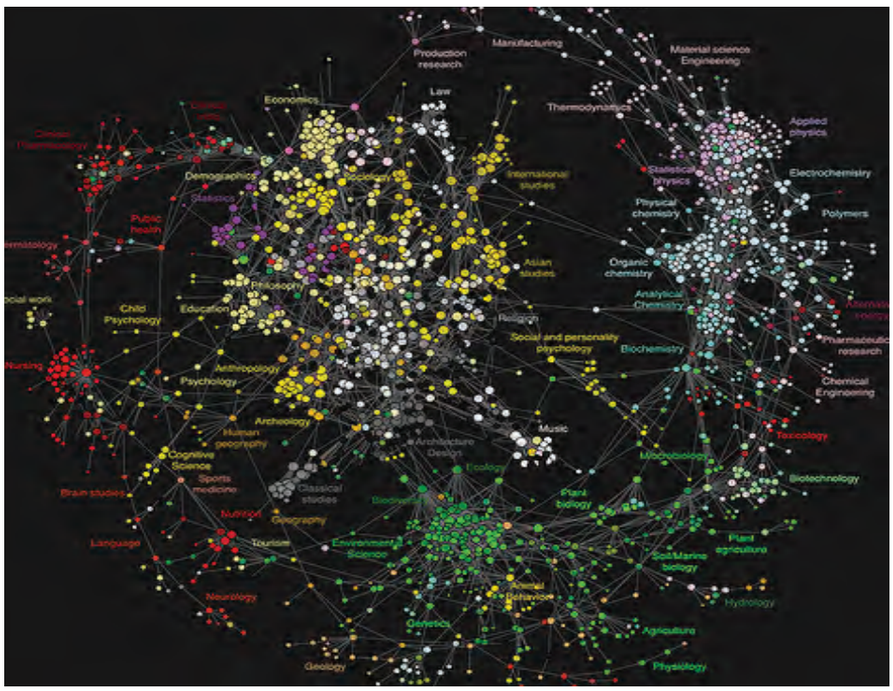



Photo credit: Bollen J, Van de Sompel H, Hagberg A, Bettencourt L, Chute R, et al. (2009) Clickstream Data Yields High-Resolution Maps of Science. PLoS ONE 4(3): e4803. doi:10.1371/journal.pone.000480

**Enables replication and verification of research results.** When data are archived and shared, results are repeatable and data can be used for re-analysis, backing up original research findings. They may also be used to expose errors or inconsistencies with original data analysis.

**Enhances collaboration:** Publicly available data enable researchers to collaborate with each other by sharing data sets, research environments and tools.

[4] Piwowar, HA, Day, RS, Fridsma, DB. Sharing Detailed Research Data Is Associated with Increased Citation Rate. 2007. PLoS ONE 2(3): http://dx.doi.org/doi:10.1371/journal.pone.0000308

## The Protein Data Bank (PDB)

The Protein Data Bank (PDB) is an open access database accelerating scientific progress and discovery. Created in 1971, the PDB initially contained seven solved protein structures. Since its inception the PDB has grown rapidly, and currently contains over 43,000 solved structures. The use of the data bank has also grown rapidly- from 1 million downloads a month in early 2007, to 5 million downloads per month in 2008.

www.wwpdb.org

# The Current Context

Approaches to the management of research data vary significantly according to discipline. Some fields, such as genomics, proteomics, high-energy physics, and astronomy have long-standing traditions of data archiving and sharing. Others, such as chemistry and the humanities and social sciences have less established traditions. Researchers in some fields oppose sharing their data for a number of reasons, such as: ownership/intellectual property concerns; lack of awareness of the value of data sharing; or, lack of knowledge or time to prepare data for dissemination.

## International

There are a growing number of large discipline-based data archives like PubChem, GenBank, Protein Data Bank, Digital Sky Survey, World Data Centers (solar, geophysical and related environmental data), Global Biodiversity Information Facility, International Virtual Observatory Alliance, the Interuniversity Consortium for Political and Social Research (ICPSR), and so on. These archives are international in scope and aim to collect data from around the world. They also provide broad access to the data they collect in order to further research and knowledge creation. The vast majority of these archives are funded through national government departments and/ or funding agencies in the country which they are housed- although there are also a few international data archives maintained by private industry. In addition to these international, discipline-based archives, many countries maintain national data archives that have a mandate to collect valuable national data in a given subject area. For example most countries collect and archive population data, climate data, health data and so on.

In 2004, 34 countries, including Canada, signed the OECD "Declaration on Access to Research Data From Public Funding" [5]. The premise of the declaration is that publicly funded research data should be openly available to the maximum extent possible. Following on this, a number of countries are investigating how they can more systematically exploit the data created through publicly funded research. For example:

| | |
|---|---|
| **United States** | In the US, the National Science Foundation (NSF) announced a $100 million funding program called "Sustainable Digital Data Preservation and Access Network Partners (DataNet)" [6]. The program is developing "new methods, management structures and technologies to manage the diversity, size, and complexity of current and future data sets and data streams by creating a set of exemplar national and global data research infrastructure organizations".<br><br>Also in the US, an Interagency Working Group on Digital Data was formed in December 2006. Nearly 30 government agencies, offices, and councils were named as members or participants, reflecting the broad range of interests in digital scientific data. The group's final report proposes the following strategy for the US: "Create a comprehensive framework of transparent, evolvable, extensible policies and management and organizational structures that provide reliable, effective access to the full spectrum of public digital scientific data. Such a framework will serve as a driving force for American leadership in science and in a competitive, global information society" [7]. |
| **United Kingdom** | The UK already has a very robust infrastructure of data repositories for collecting research data, managed by several RCUK funding agencies. To build on this, the Consortium of Research Libraries in the UK and Ireland, and the IT Directors Group undertook a study to assess the feasibility and costs of developing and maintaining a national shared digital research data service for the UK Higher Education sector. The study concluded that considerable efforts are being made in other countries to surface and exploit data on a national or regional scale and that there is a need for a UK-wide approach to research data management to ensure that it does not fall behind [8]. |

[5] "Science, Technology and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 - Final Communique". 2004. Organisation for Economic Cooperation and Development. http://www.oecd.org/document/0,2340,en_2649_34487_25998799_1_1_1_1,00.html

[6] "Community-based Data Interoperability Networks". 2009. National Science Foundation, www.nsf.gov/pubs/2007/nsf07565/nsf07565.htm

[7] "Harnessing the Power of Digital Data for Science and Society". Report of the Interagency Working Group on Digital Data to the Committee on Science of the National Science and Technology Council. January 2009. pg.2 www.nitrd.gov/about/Harnessing_Power_Web.pdf

[8] "UK Research Data Service". 2009. www.ukrds.ac.uk/

**Australia**

In September 2008, representatives from the major Australian research community endorsed the Brisbane declaration, which states among other things that "Every citizen should have free open access to publicly funded research, data and knowledge" and "Every Australian university should have access to a digital repository" [9].

The Australian government recently announced the creation of the Australian National Data Service (ANDS) . ANDS began in September 2008 and was established as part of the National Collaborative Research Infrastructure Strategy. Among other stated goals, ANDS aims "to ensure that Australian research data are well managed, made available for access, and discoverable" [10].

[9] "Open Access and Research Conference 2008". 2008 Queensland University of Technology. www.oar2008.qut.edu.au/
[10] "Australian National Data Service". 2009. Monash University. http://ands.org.au/

## Canada

In October 2000, the Social Sciences and Humanities Research Council and the National Archivist of Canada established a Working Group of research and archival experts and asked them to assess the need for a national research data access, preservation and management system. The Working Group found large gaps in the infrastructure for collecting and preserving research data. The final report, published in 2002, recommended the creation of a new national research data archival service.[11]

In November 2004, the National Research Council, in partnership with Canada Foundation for Innovation, Canadian Institutes of Health Research, and the Natural Sciences and Engineering Research Council, undertook a consultation on access to research data in Canada (referred to as the National Consultation on Access to Scientific and Research Data or NCASRD). The consultation received input from the natural and medical sciences community. The final report was released in January 2005 and provides a "road map" for the implementation of a national plan for open access to publicly funded scientific research data. The recommendations in the report have not yet been acted on.

In January 2004, Canada was a signatory of the OECD Declaration on Access to Research Data From Public Funding. But, unlike some other national signatories, Canada has not responded with concrete actions to give its agreement to the MOU substance.

[11] "Final Report: National Data Archive Consultation". June 2002. Social Sciences and Humanities Research Council. www.sshrc.ca/site/about-crsh/publications/da_finalreport_e.pdf

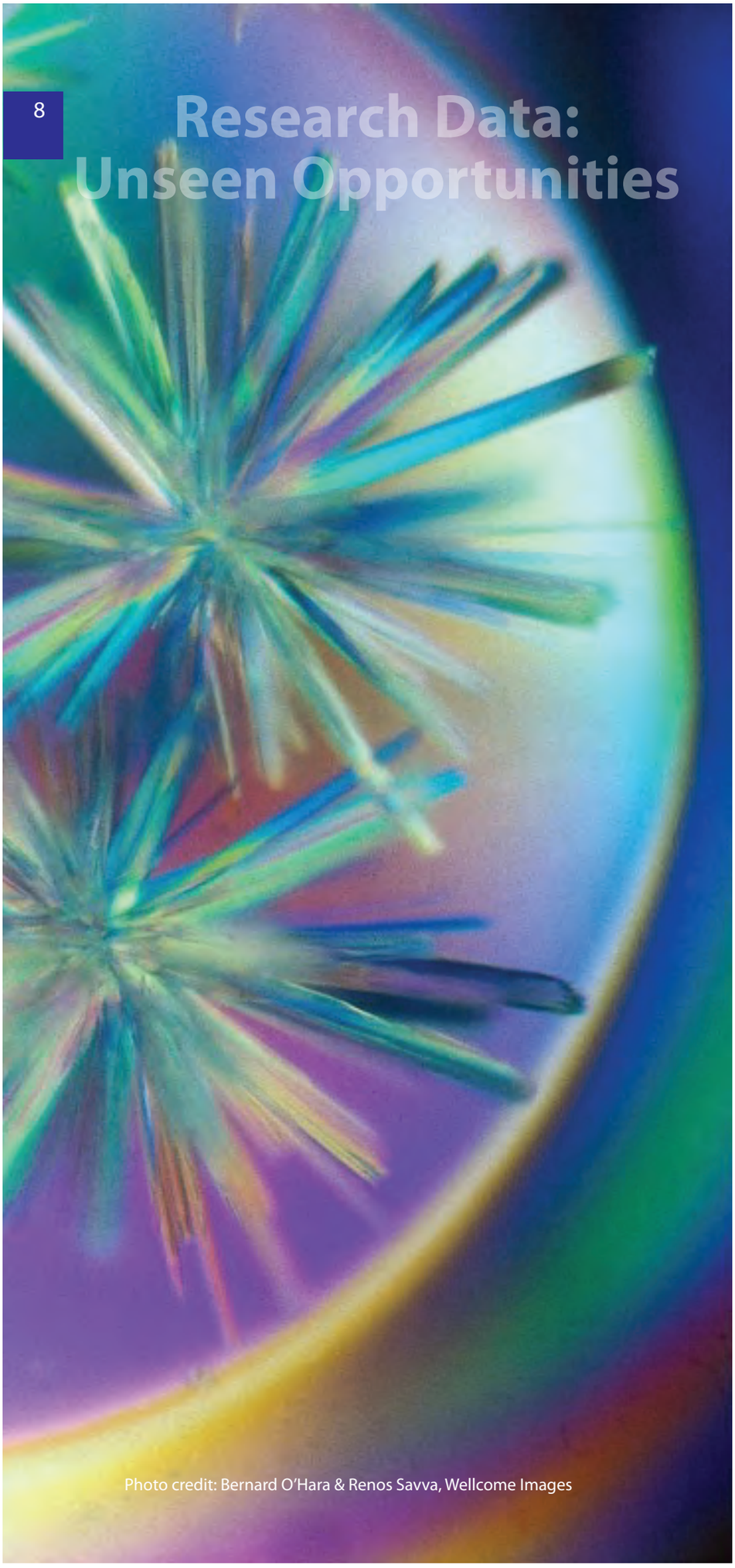



Photo credit: Bernard O'Hara & Renos Savva, Wellcome Images

# Canadian Groups Concerned with Data Management

## Canadian Research Data Strategy Working Group

On January 25, 2008, the Canada Institute for Scientific and Technical Information (CISTI) hosted a meeting with key stakeholders to discuss how to begin implementing the National Consultation on Access to Scientific and Research Data outcomes[12]. A Working Group was established with membership from all the federal funding agencies, Canadian Association of Research Libraries (CARL), Compute Canada, CANARIE, and individual researchers. The group published a Gap Analysis in fall 2008 and is currently working to address the gaps to the stewardship of research data in Canada.

[12] "Research Data Canada". 2009. Canada Institute for Scientific and Technical Information. http://data-donnees.gc.ca/

## Canadian National Committee for CODATA

(CNC/CODATA)

This committee is the Canadian voice of CODATA (Committee on Data for Science and Technology)[13]. CODATA is an interdisciplinary Scientific Committee of the International Council for Science (ICSU), which was established 40 years ago. CNC/CODATA is sponsored by CISTI (Canada Institute for Scientific and Technical Information), part of the National Research Council of Canada. CODATA has no formal governmental status, but works through its national members and task groups to "improve the quality, reliability, management and accessibility of data of importance to all fields of science and technology".

[13] "Canadian National Committee for CODATA". 2009. www.codata.org/canada

# Case Studies of Data Management Activities in Canada

## Ontario Data Documentation, Extraction Service and Infrastructure Initiative

ODESI is a jointly funded project between the Ontario Council of University Libraries (OCUL) and OntarioBuys that provides university researchers with unprecedented access to a significant number of datasets. The project targets Statistics Canada datasets,[14] data files from Gallup Canada and other polling companies, public-domain files such as the Canadian National Election Surveys and selected files from the Inter-University Consortium for Political and Social Research (ICPSR). The files are marked-up using Data Documentation Initiative (DDI), an international, XML-based metadata tagging system which allows data resource discovery, distributed access, extraction and analysis.

[14] "Ontario Data Documentation, Extraction Service and Infrastructure". 2009. Ontario Council of University Libraries. http://search2.odesi.ca


## International Polar Year

The International Polar Year is a large scientific program focused on the Arctic and the Antarctic from March 2007 to March 2008. IPY has a comprehensive data policy which "requires that IPY data, including operational data delivered in real time, are made available fully, freely, openly, and on the shortest feasible timescale."[15] Dozens of Canadian research projects were selected for IPY funding from a variety of sources including the federal government, territorial governments, granting agencies and foundations. Each project was required to develop a Data Management Plan in order to receive funding. It is likely that many Canadian projects will have difficulty adhering to the IPY policy, as there are not a lot of repositories available in Canada to collect, preserve and provide access to the data they create.

[15] "International Polar Year 2007-2008 Data Policy". April 2008. http://classic.ipy.org/Subcommittees/final_ipy_data_policy.pdf

# Case Studies of Data Management Activities in Canada

## VENUS

The Victoria Experimental Network Under the Sea (VENUS),[16] led by the University of Victoria, is a cabled ocean observatory. VENUS delivers real time information from the seafloor via fibre optic cables connected to instruments at the University of Victoria, BC, where they are archived. The VENUS portal gives scientists and the general public free access to a constantly expanding "warehouse" of images, sounds and live data from the ocean floor[17] to facilitate ocean research anywhere in the world.

[16, 17] "VENUS: Victoria Network Under the Sea". 2009. University of Victoria. www.venus.uvic.ca

## Canadian Barcode of Life

The Canadian Barcode of Life Network, along with the large international consortium of which it is a part, are developing an accurate, rapid, cost-effective and universally accessible DNA-based system for species identification. DNA barcoding is a highly promising approach to resolve the 'taxonomic impediment' that constrains much biodiversity research. All these barcoding projects share the goal of building an open-access database of reference barcodes that will improve our understanding of biodiversity and will allow non-taxonomists to identify species. As of August 2009, researchers had deposited over 670,000 barcode records from 61,000 species in the international Barcode of Life Database (the vast majority–close to 500,000-come from the Canadian Centre.[18]

[18] "Barcode of Life Data Systems". 2009. Biodiversity Institute of Ontario. www.barcodinglife.org/views/login.php

## Research Data Centres Metadata Project

This project will produce a set of tools that will enable researchers to provide enhanced metadata for the entire life-cycle of their research studies, using the DDI 3 XML schema. The project, based out of the University of Manitoba, will mark up files in the Research Data Centres and create tools for migration to Data Documentation Initiative 3. It will produce a set of tools that will enable researchers to provide enhanced metadata for the entire life-cycle of their projects. A Research Metadata Centre has been established in Gatineau, Quebec to produce compliant metadata for Statistics Canada's confidential files in Canada's Network of Research Data Centres and an international software consortium headed by Brechenhill in Canada is creating tools to migrate data documentation to the Data Documentation Initiative 3.

# Gaps in Data Stewardship in Canada



The Research Data Strategy Working Group lead by CISTI conducted a Gap Analysis of the stewardship of research data in the fall of 2008.[19] The analysis revealed significant barriers to access and preservation of research data — barriers that, if not addressed, could have a serious impact on the future of Canadian research and innovation. The gaps identified are:

**Data policies:** Data policies in Canada do not cover all types of research data; and in some cases, researchers are not adhering to policies. Policies addressing the preservation of data are lacking.

**Funding for data management:** The funding structure for research in Canada does not support data management beyond the life of a given research project. There are not many institutions in Canada that provide sustainable funding for data repositories. As well, there are significant costs associated with preparing data for dissemination and these costs are not supported through existing funding mechanisms (i.e. grants or university budgets).

**Data repositories:** Only a few active data repositories in Canada allow researchers to deposit their data. There are large gaps in both coverage and capacity of data repositories. Repositories do not exist for all subject areas, and the vast majority of research data still resides on researchers' hard drives.

**Skills:** Researchers rarely have the skills required to appropriately manage their data. And, there are few data scientists or information professionals with knowledge of data cataloguing, metadata standards and processes, preservation management and assessment of the value of data to support researchers.

**Standards:** Not all researchers and disciplines have adopted international standards. In terms of interoperability, there is a lot of work to be done to enable interoperability of data from different projects and across disciplines. Metadata standards, which are particularly important in ensuring that data are accessible and interoperable with other types of data sets, are not always adhered to by data creators/ collectors.

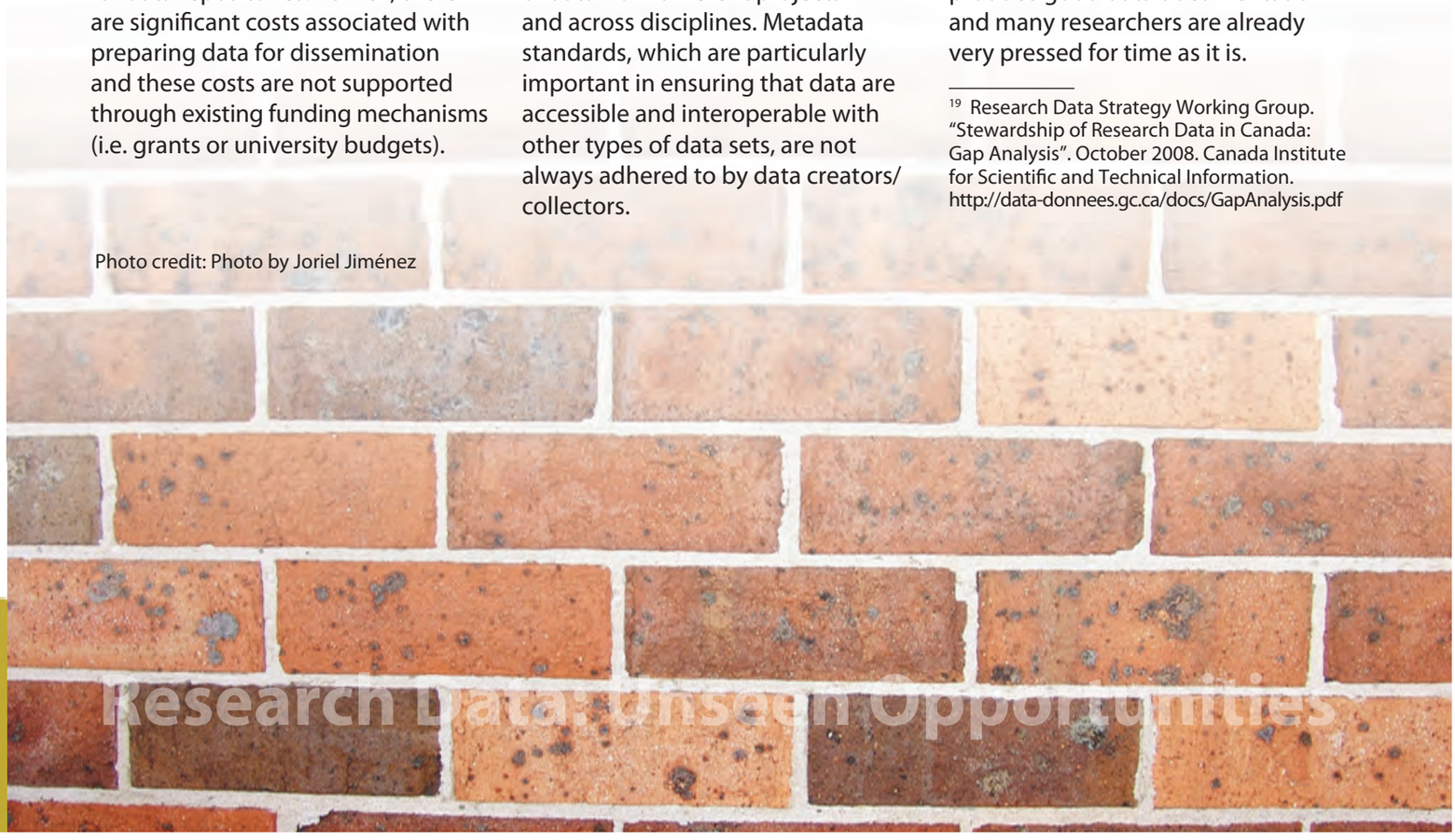
**Incentives:** There are very few incentives for researchers to share data. There are also real disincentives to sharing data, because researchers are pressed to get as many publications from their primary data collections as possible before sharing the data.

**Roles and Responsibilities:** With the exception of some government departments, there are no national institutions responsible for preserving, managing and making research data publicly accessible on the scale required that supports the needs of stakeholder communities.

**Time:** A crucial part of creating a good dataset with long-lasting usability is ensuring that the data are easy to understand and analyze. This requires accompanying data description and documentation that is user-friendly, clear and detailed, yet comprehensive. It takes time to produce good data documentation and many researchers are already very pressed for time as it is.

[19] Research Data Strategy Working Group. "Stewardship of Research Data in Canada: Gap Analysis". October 2008. Canada Institute for Scientific and Technical Information. http://data-donnees.gc.ca/docs/GapAnalysis.pdf

Photo credit: Photo by Joriel Jiménez



# Data Management Policies in Canada

A mixture of policies issued by federal/provincial governments, funding agencies, institutions, and even some research projects governs data management activities in Canada. The policies vary considerably and in some cases, cover different aspects of data management. However, most researchers are required, at a minimum, either by their institution or their funding agency (or both), to retain their research data up to 5 years after the completion of a research project.

## Funding Agency Policies

| | |
|---|---|
| **CIHR** | Canadian Institutes of Health Research requires grant recipients to deposit certain data types- bioinformatics, atomic, and molecular coordinate data- into the appropriate public database immediately upon publication of research results. CIHR also requires researchers to retain original data sets arising from CIHR-funded research for a minimum of five years after the end of the grant. They have plans to review this policy in the near future and possibly expand it. [20] |
| **SSHRC** | Social Sciences and Humanities Research Council has had a Research Data Archiving Policy in place since 1990. The policy requires that "All research data collected with the use of SSHRC funds must be preserved and made available for use by others within a reasonable period of time. SSHRC considers "a reasonable period" to be within two years of the completion of the research project for which the data was collected." [21] SSHRC has not been enforcing this policy with researchers and fewer than 10% of researchers have adhered to the conditions of the policy since its inception.[22] |
| **NSERC** | Natural Sciences and Engineering Research Council has no general policy in regards to research data. They do, however, have guidelines for researchers funded through the NSERC Strategic Networks Program. The guidelines state, "To encourage the sharing and dissemination of research data and its use by others within a reasonable period of time, an agreement regarding responsibility for the maintenance and preservation of large data sets must be in place at the outset of network activities." [23] |
| **All three** | Tri-Council Policy Statement on the Ethical Conduct for Research Involving Humans<br><br>"Researchers should ensure that the data obtained are stored with all the precautions appropriate to the sensitivity of the data. Data released should not contain names, initials or other identifying information. While it may be important to preserve certain types of identifiers (e.g., region of residence), these should be masked as much as possible using a standardized protocol before the data are released for research purposes."[24] |

[20] "Policy on Access to Research Outputs". September 2007. Canadian Institutes of Health Research. www.cihr-irsc.gc.ca/e/34846.html
[21] "SSHRC Research Data Archiving Policy". 2009. Social Sciences and Humanities Research Council. http://www.sshrc.ca/site/apply-demande/policies-politiques/edata-donnees_electroniques-eng.aspx
[22] Humphrey, C. "Preserving research data: A time for action". In Preservation of electronic records: new knowledge and decision-making. 2003. Canadian Conservation Institute: Ottawa. pg. 83-89.
[23] "Strategic Network Grants Program". 2009. Natural Sciences and Engineering Research Council. www.nserc-crsng.gc.ca/Professors-Professeurs/RPP-PP/SNG-SRS_eng.asp
[24] "Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans: Draft 2nd Edition". December 2008. Panel on Research Ethics. http://pre.ethics.gc.ca/eng/policy-politique/initiatives/draft-preliminaire/

## Other Data Policies

| | |
|---|---|
| **McGill University**<br><br>(No other university policies were found in putting together this document) | "Data must be organized in a manner that allows ready verification….Subject to exceptions based on a duty of confidentiality and the laws respecting intellectual property and access to information, after data are published, they must be made available to any party presenting a reasonable request to examine them. In cases where there is a disagreement between the researcher and the person requesting the data, the matter shall be referred to the Office of the Vice Principal Research for resolution…(a)ll original data must be retained for a reasonable length of time. A period of at least five years from the date of publication is recommended."[25] |
| **Fisheries and Oceans Canada**<br><br>**(One example of a government department data policy)** | "Management Policy for Scientific Data" states, "It is the responsibility of Science and Oceans managers to ensure that data collectors under their management submit their data as well as data collected under contract to or partnership with other agencies, to the appropriate data centre in a timely fashion. This is important to ensure that data are quickly migrated into a 'managed' environment where they are properly backed up and secured from accidental or circumstantial loss, and where the supporting metadata are integrated with the data to preserve the long-term usefulness of a data set."[26] |
| **International Polar Year (IPY)** | The "IPY Data Policy" states, "in order to maximize the benefit of data gathered under the auspices of the IPY, the IPY Joint Committee requires that IPY data, including operational data delivered in real time, are made available fully, freely, openly, and on the shortest feasible timescale"[27]. In addition, "All IPY data must be archived in their simplest, useful form and be accompanied by a complete metadata description."[28] |
| **Federal Government: Personal Information Protection and Electronic Documents Act (PIPEDA)** | All data in Canada are also subject to the Personal Information Protection and Electronic Documents Act.<br><br>Personal information shall not be used or disclosed for purposes other than those for which it was collected, except with the consent of the individual or as required by law. Personal information shall be retained only as long as necessary for the fulfillment of those purposes. Personal information that is no longer required to fulfill the identified purposes should be destroyed, erased, or made anonymous. Organizations shall develop guidelines and implement procedures to govern the destruction of personal information.[29] |

[25] "Policy on Research Ethics". April 7, 2008. McGill University. www.mcgill.ca/files/secretariat/Research-Ethics-Policy-on.pdf
[26] "Management Policy of DFO Scientific Data". June 12, 2001. Fisheries and Oceans Canada. http://www.osl.gc.ca/sgdo/en/info-donnees/acces-donnees.html
[27, 28] "International Polar Year 2007-2008 Data Policy". April 2008. http://classic.ipy.org/Subcommittees/final_ipy_data_policy.pdf
[29] "The Personal Information Protection and Electronic Documents Act". 2009. Department of Justice Canada. http://laws.justice.gc.ca/en/P-8.6/FullText.html

# Responses to Concerns about Data Management

## I can't archive and share my data because they come from human subjects.

Respect for privacy in research is an internationally recognized norm and ethical standard. In Canada, the three major granting councils have developed a shared policy statement to govern all research involving human participants undertaken in Canada. The policy states that personal information of study participants should not be shared. However, the policy also recommends that:

"As a general rule, the best protection of the confidentiality of personal information and records will be achieved through anonymity. If the data being stored are truly anonymous, the research project will need only minimal REB [Research Ethics Board] scrutiny."[30] (Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans)

There are well-established procedures for anonymizing data and researchers can consult with the local data professional at their research institution about whether and how to anonymize data containing personal information so that it can be shared with others. There is some work to be done in terms of raising awareness of anonymization techniques with Research Ethics Boards (REBs). Despite Tri-council advice to anonymize data in the case of human subjects, many REBs still often ask researchers to destroy data.

## Sound data management practices are too costly.

While there are significant costs involved in data management such as the application of metadata, maintenance of repositories, and data preservation activities- the benefits far outweigh the costs. Re-creating research data sets can be prohibitively expensive. In many cases it is impossible to re-create lost data, as researchers, test subjects, testing conditions, and so on disperse or disappear over time. In these circumstances, maintaining a reliable, managed environment for protecting the considerable institutional investment involved in creating research data would represent a comparatively small cost when placed against the prospect of the higher and perhaps prohibitive costs of re-creation later on or the complete and irretrievable loss of data.

[30] "Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans: Draft 2nd Edition". December 2008. Panel on Research Ethics. http://pre.ethics.gc.ca/eng/policy-politique/initiatives/draft-preliminaire/

## If I share my data, I won't be able to fully capitalize on any possible patents or other economic benefits.

Sound data management principles are not in contradiction with intellectual property and data ownership. Investigators can choose to restrict access to data to maximize their professional and economic benefit, by postponing the sharing of data until publication or application of patent, or by applying a non-commercial use license to the data.

The OECD advises that consideration should be given to measures that promote non-commercial access and use while protecting commercial interests, such as delayed or partial release of such data, or the voluntary adoption of licensing mechanisms. Such measures can allow the primary participants to fully exploit the research data without unnecessarily shutting off access.

Photo credit: http://www.apputechnologies.com

# What can be done on campus?

## Researchers

Commit to sharing research data as openly as possible.
Develop data management plans before the beginning of a research project.
Understand and comply with funding agency data management policies.
Make use of the data professionals on campus to assist in collecting and managing research data.
Identify and use international standards for data management.

## University Administrators

Develop policies that support sound data management activities.
Support the implementation of data repositories at the institution.
Provide education for researchers about data management practices.
Provide support for researchers by hiring qualified data scientists or librarians and make these professionals known to (and part of) the appropriate research teams.
Recognize data sharing contributions in hiring, and promotion and tenure decisions.

## Research Libraries

Develop and manage data repositories at the institution.
Support training for librarians in the area of data stewardship.
Provide support for researchers by hiring qualified data librarians and make these professionals available to the research community.
Provide education for researchers about data management practices.


**65 University Private, Room 238, Ottawa, ON K1N 9A5**
**Tel: 613-562-5385 • Fax: 613-562-5297**
**carladm@uottawa.ca • www.carl-abrc.ca**

The Canadian Association of Research Libraries (CARL) was established in 1976 and consists of Canada's 28 major academic libraries plus Library and Archives Canada, the Canada Institute for Scientific and Technical Information (CISTI) and the Library of Parliament. Membership is institutional and is open primarily to libraries of Canadian universities which have PhD programs in the arts, social sciences and sciences.

CARL strives to enhance the capacity of Canada's research libraries to partner in research and higher education, seeking effective and sustainable scholarly communication and public policy encouraging of research and broad access to scholarly information.

CARL is an associate member of the Association of Universities and Colleges of Canada (AUCC), and is incorporated as a non-profit organization under the Canada Corporations Act.

**CARL Member Institutions**

University of Alberta
University of British Columbia
Brock University
University of Calgary
Carleton University
CISTI (Canada Institute for Scientific and Technical Information)
Concordia University
Dalhousie University
University of Guelph
Université Laval
Library and Archives Canada
Library of Parliament
University of Manitoba
McGill University
McMaster University
Memorial University of Newfoundland
Université de Montréal
University of New Brunswick
University of Ottawa
Université du Québec à Montréal
Queen's University
University of Regina
University of Saskatchewan
Université de Sherbrooke
Simon Fraser University
University of Toronto
University of Victoria
University of Waterloo
University of Western Ontario
University of Windsor
York University

Executive Director: **Brent Roe**

President: **Ernie Ingles**, FRSC (2009 – 2011)