In the last five years, the world produced more research data than has been created or lost because it is not being systematically managed. 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.

In 2004, 34 countries, including Canada, signed the OECD ‘Declaration on Access to Research Data from Public Funding’.

The promise of the declaration is that publicly funded research should be openly available to the maximum extent possible. The aims of the OECD Declaration are to give its commitment to the OECD Declaration substance.

Following on this, a number of countries have begun to investigate a coordinated and national approach to managing research data in that country. The aims of the National Data Service (ANDS) in order to improve the sharing and reuse of research data.

Canada, however, has not yet responded with concrete actions to give its commitment to the OECD Declaration substance.

Both the US and the UK governments have recently launched open data initiatives that increase public access to research data. The governments are trying to fulfill these requirements.

Major Benefits of Data Management, Sharing, and Reuse

Accelerates scientific progress

Data sharing allows researchers to access and understand others’ data and re-use them for their own scientific purposes, thereby speeding up the rate of new discoveries.

A coordinated and national approach to managing research data in Canada is required in order to derive greater and longer term benefits, both socially and economically, from the extensive public investments that are made in research.

The Victoria Experimental Network Under the Sea (VENUS) project, supported by the Canadian Federal and Provincial Governments, successfully delivered real-time data from the subsea wave to fibre optic cables connected to instruments at the University of Victoria where they are archived. VENUS researchers have spent over half a trillion measurements back to the University of Victoria and, through the Internet, to scientists at the rest of the world. To handle the huge volume of data, the University of Victoria has developed an advanced data centre and an archiving system that allows the raw and processed data to be made available for at least the next 25 years to everyone - both the public and academic communities - in Canada.

In 2007, the Journal of Applied Developmental Psychology published a special issue to highlight how the Study of Early Child Care and Youth had a number of new initiatives that had been used by researchers to address a range of research questions not envisioned in the original study plan. The data sets contain longitudinal data about the family, day care and school environments of over 1000 children, tracked from birth through age fifteen. Most of the publications by original study investigators focused on child-care linkages. However, editors found that the original data had been used by hundreds of other investigators exploring a wide range of other research problems, such as child development, women’s employment patterns, quality of academic instruction, family social support, and psychosocial processes, child-adult interaction, and peer interaction.

Current Context

Approaches to the management of research data vary significantly according to discipline. Some fields, such as genetics, 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.

The International Polar Year (IPY) was the largest-ever international program of scientific research focused on the Arctic and Antarctic regions. It involved thousands of researchers from over 60 nations and produced large amounts of diverse research data. IPY researchers are required to share their data with others in a timely manner and effectively manage their research data to ensure the data are available for future research. The lack of available data repositories is posing a challenge for many IPY researchers who are struggling to fulfill these requirements.

Research data: unseen opportunities

Data sharing allows researchers to access and understand others’ data and re-use them for their own scientific purposes, thereby speeding up the rate of new discoveries.

A coordinated and national approach to managing research data in that country. The aims of the National Data Service (ANDS) in order to improve the sharing and reuse of research data.

Canada, however, has not yet responded with concrete actions to give its commitment to the OECD Declaration substance.

Following on this, a number of countries have begun to investigate a coordinated and national approach to managing research data in that country. The aims of the National Data Service (ANDS) in order to improve the sharing and reuse of research data.

Canada, however, has not yet responded with concrete actions to give its commitment to the OECD Declaration substance.

Both the US and the UK governments have recently launched open data initiatives that increase public access to research data. The governments are trying to fulfill these requirements.

Major Benefits of Data Management, Sharing, and Reuse

Accelerates scientific progress

Data sharing allows researchers to access and understand others’ data and re-use them for their own scientific purposes, thereby speeding up the rate of new discoveries.

A coordinated and national approach to managing research data in Canada is required in order to derive greater and longer term benefits, both socially and economically, from the extensive public investments that are made in research.

The Victoria Experimental Network Under the Sea (VENUS) project, supported by the Canadian Federal and Provincial Governments, successfully delivered real-time data from the subsea wave to fibre optic cables connected to instruments at the University of Victoria where they are archived. VENUS researchers have spent over half a trillion measurements back to the University of Victoria and, through the Internet, to scientists at the rest of the world. To handle the huge volume of data, the University of Victoria has developed an advanced data centre and an archiving system that allows the raw and processed data to be made available for at least the next 25 years to everyone - both the public and academic communities - in Canada.

In 2007, the Journal of Applied Developmental Psychology published a special issue to highlight how the Study of Early Child Care and Youth had a number of new initiatives that had been used by researchers to address a range of research questions not envisioned in the original study plan. The data sets contain longitudinal data about the family, day care and school environments of over 1000 children, tracked from birth through age fifteen. Most of the publications by original study investigators focused on child-care linkages. However, editors found that the original data had been used by hundreds of other investigators exploring a wide range of other research problems, such as child development, women’s employment patterns, quality of academic instruction, family social support, and psychosocial processes, child-adult interaction, and peer interaction.

Current Context

Approaches to the management of research data vary significantly according to discipline. Some fields, such as genetics, 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.

The International Polar Year (IPY) was the largest-ever international program of scientific research focused on the Arctic and Antarctic regions. It involved thousands of researchers from over 60 nations and produced large amounts of diverse research data. IPY researchers are required to share their data with others in a timely manner and effectively manage their research data to ensure the data are available for future research. The lack of available data repositories is posing a challenge for many IPY researchers who are struggling to fulfill these requirements.

End Notes

1. According to Tony Hey, former Director of the e-Science Care programme, in his talk at the British Library Conference Centre, London UK, Wednesday 23 October 2008.
2. From: http://www.oecd.org/document/0,2340,39456655_33341582,00.html
5. From: http://www.w3.org/2002/08/ov/010

This brochure was created in June 2010 by Kahlen Schayer on behalf of the Canadian Association of Research Libraries Data Management Subcommittee.

Canadian Association of Research Libraries
Montréal 1400 Boulevard Edouard-Montpetit Suite 216 Montreal, Quebec Canada
K1B 5A9
Phone: 613.562.5264
Fax: 613.562.5297
Email: carladm@uottawa.ca
www.ca-rldc.net

© 2010 CARL, The content in this document is licensed under the Creative Commons Attribution-NonCommercial-2.5 Canada License.
Gaps in Data Stewardship in Canada

The Research Data Strategy Working Group, a multidisciplinary group led by CISTI (Canada Institute for Scientific and Technical Information) conducted a gap analysis of the stewardship of research data in Canada in the fall of 2008. The gaps reported were as follows:

Data policies: Data policies in Canada do not cover all types of research data; and in some cases, researchers do not adhere to policies.

Funding for data management: The funding structure for research in Canada does not support data management beyond the life of a given research project.

Data repositories: Only a few active data repositories in Canada allow researchers to deposit their data.

Skills: Researchers rarely have the skills required to appropriately manage their data and there are few data professionals to assist them.

Standards: Not all researchers and disciplines have adopted international standards.

Incentives: There are very few incentives for researchers to share 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.

Time: It takes time to produce good data documentation and many researchers are already very pressed for time as it is.

The lack of incentives for researchers is regarded as a crucial unresolved obstacle to establishing a data sharing culture. In order to address this, the research community needs to develop and test incentives for data management sharing. Data sharing contributions could then be considered during hiring, tenure, and promotion decisions by institutions.

Relevant Research Data Policies in Canada - Funding Agencies

Canadian Institutes of Health Research
Researchers are already very pressed for time as it is.

Time:

departments, there are no national institutions responsible for

Roles and Responsibilities:

There are very few incentives for

international standards.

Standards: Not all researchers and disciplines have adopted

assist them.

Natural Sciences and Engineering Research Council
Natural Sciences and Engineering Research Council

No general policy in regards to data sharing.

Tri-Council Policy Statement on the Ethical Conduct for

Research Involving Humans states that the "best protection of the confidentiality of personal information and records will be achieved through anonymity if the data being shared are truly anonymous; if the research project will need only minimal REB [Research Ethics Board] scrutiny."

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.

Sound data management practices are too costly.

Maintaining a reliable, managed environment for protecting the considerable investment involved in creating research data represents a comparatively small cost when placed against the prospect of the higher and perhaps prohibitive costs of re-creating data later on or the complete and irretrievable loss of data.

Examples of Other Data Policies

Fisheries and Oceans Canada

"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 partnerships with other agencies, to the appropriate data centres in a timely fashion."

International Polar Year (IPY)

"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 freely, openly, and on the shortest feasible timelines."

Responses to Concerns about Data Management

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.

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

The Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans states that the "best protection of the confidentiality of personal information and records will be achieved through anonymity if the data being shared are truly anonymous; if the research project will need only minimal REB [Research Ethics Board] scrutiny."

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.

Sound data management practices are too costly.

Maintaining a reliable, managed environment for protecting the considerable investment involved in creating research data represents a comparatively small cost when placed against the prospect of the higher and perhaps prohibitive costs of re-creating data later on or the complete and irretrievable loss of data.

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 available to the appropriate research teams.

Recognize data sharing contributions in hiring, 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.

Research Data: Unseen Opportunities

An awareness brochure produced by the Canadian Association of Research Libraries (CARL)