The 2011 Canadian
Research Data Summit

Mapping the Data Landscape:

Report of the 2011 Canadian Research Data Summit

December 2011

Produced by the Research Data Strategy Working Group

Contents

I. Foreword

Canada invests billions of dollars in research each year, research that produces huge amounts of data. If properly managed, these data hold virtually unlimited potential to be re-used in innovative ways – by industry, policy makers, researchers and citizens. Unfortunately, in Canada, this potential remains unrealized. Canada is one of the few advanced countries that do not yet have a national plan for managing the research data produced through public funding. As a result, valuable data are under-utilized and an important publicly funded asset is being wasted.

The way that we, as a nation, choose to manage our research data will directly impact our ability to undertake leading-edge research and development in the future. But, managing data is about much more than supporting research excellence. Digital data are the raw materials of the knowledge economy, and are becoming increasingly important for all areas of society, including industry. A recent report by McKinsey and Company asserts, "Like other essential factors of production such as hard assets and human capital, it is increasingly the case that much of modern economic activity, innovation and growth simply couldn't take place without data."[1] In addition, evidence-based policies and practices, which have demonstrably better outcomes, cannot be developed without the availability of relevant data.

Urgent action is needed. Canada's lack of a strategy and action plan places us at a disadvantage compared to our international competitors. A coordinated and national approach to managing and providing access to research data is required to ensure that Canadians and others derive greater and more long-term benefit, both socially and economically, from the extensive public investments that are made in research. Such an approach would represent a logical extension of the Government of Canada's Open Government Initiative, which already aims to make government generated data widely available.

The [Research Data Strategy Working Group](#) (RDSWG) is a collaborative effort launched in 2008. It is a multi-disciplinary group of universities, institutes, libraries, operators of research infrastructure, granting agencies, governments, and individual researchers that are united through a shared recognition of the pressing need to deal with Canadian data stewardship issues. The RDS Working Group's activities focus on the actions and leadership roles that researchers, institutions and governments must take to ensure Canada's research data are accessible and usable for current and future generations.

In September 2011, the RDSWG hosted *Mapping the Data Landscape: the 2011 Canadian Research Data Summit* to develop a national approach for managing Canada's research data. Based on input from Summit participants, the RDSWG has developed a draft National Strategy for Research Data in Canada, presented in this report. The National Strategy, which contains a vision, goals and framework for action, will be made available online in early 2012 for further input and endorsement.

---

[1] McKinsey & Company, [Big Data: The next frontier for innovation, competition, and productivity,](#) March 2011

II. About Research Data

Research data are defined here as the "factual records used as primary sources for research, and that are commonly accepted in the research community as necessary to validate research findings."[2]

The concept of research data is complex and fluid. Virtually all types of digital information have the potential to be research data if they are being used as a primary resource for research. The Australian National Data Service describes the wide variety of data types and forms as follows: "Some of the data might be raw data, the unprocessed observations of particular phenomena. Some might be processed data, the data produced when raw data has been calibrated or corrected. Some might be derived data, which present a summary or specific view of the raw data. Some might be textual data, the publications which result from a research project or the textual data (texts, bibliographies, surveys, etc.) which forms the basis of a research project."[3]

All of these types of data, in their many forms and sizes are encompassed in the term 'research data' as used in this report. Data management policies and practices must take into account the diversity of data types, working methods, curation practices, and needs across the various disciplines and specializations.

---

[2] OECD Principles and Guidelines for Access to Research Data from Public Funding. OECD, 2007. Available at: www.oecd.org/dataoecd/9/61/38500813.pdf

[3] What is research data? ANDS Guide. Australian National Data Service. Available at: ands.org.au/guides/research-data-australia.pdf

1. Introduction

> *"Research cannot flourish if data are not preserved and made accessible.*
> *All concerned must act accordingly."* (*Nature*, September 9, 2009)

Digital data are revolutionizing the way research is being carried out, leading to a new data-centric way of thinking. However, as the volume of research data grows exponentially, so must the efforts to ensure that they are preserved, accessible, and understandable. For research data to be available for future use, long-term preservation must be a goal at the time data are created.

On September 14 and 15, 2011, the Research Data Strategy Working Group (RDSWG) hosted *Mapping the Data Landscape: The 2011 Canadian Research Data Summit.* The Summit brought together over 150 senior researchers, high-level policy makers, university administrators, and members of the private sector. Together, participants worked on formulating a shared strategy for addressing the challenges and opportunities for maximizing the benefits of our collective investment in research data in Canada.

World-class speakers set the stage for an engaging and productive day. The opening keynote featured Dr. John Wood, Secretary General of the Association of Commonwealth Universities and Chair of the European Union High Level Group on Scientific Data. Other speakers and moderators included Corinne Charest, Chief Information Officer of the Government of Canada; Dr. Chad Gaffield, President of the Social Sciences and Humanities Research Council; Dr. David H. Turpin, President and Vice-Chancellor of the University of Victoria; and Dr. Tony Hey, Corporate Vice President, Microsoft Research.

The Summit program was interspersed with presentations from Canadian researchers representing a wide variety of disciplines. Presenters spoke about how they use data in their research and the many exciting opportunities for new discovery now that data are in digital format. From brain mapping, to detecting Pulsars with radio telescopes, to humanities computing, virtually all fields of research are becoming highly data dependent, and as such will require support for data management activities. Yet, in Canada, there are barriers that must be overcome if we are to take full advantage of the digital data revolution. Even in fields that are already well-advanced in this area, such as genomics, there remain challenges. These include a lack of long-term preservation; the need for more analytical tools; and an absence of skilled workers for managing research data. In addition, projects that use aggregated data from disparate sources, like the Canadian Index of Wellbeing and international biodiversity research, are hampered because the data they need are often not available at all, or they do not adhere to appropriate standards. The inherent tensions between intellectual property and confidentiality versus data sharing are also recognized issues; as is the pervasiveness of a culture of "ownership" in the research milieu.

Despite the many challenges, there was widespread consensus at the Summit around the tremendous potential value for research data and recognition that there is a need to work together to address these challenges. Summit participants discussed ways of overcoming the barriers, and identified goals and actions that would constitute a national strategy for research data in Canada. The variety of perspectives and depth of knowledge of attendees made for a

comprehensive and extremely worthwhile discussion that serves well as the foundation for a made-in-Canada approach for maximizing the availability of research data.

This report aims to capture the major points discussed at the Summit; proposes the first draft a National Strategy for Research Data in Canada; and presents the critical next steps for the Research Data Strategy Working Group in order to move forward. The RDS Working Group will seek endorsement and further input for the National Strategy from the broader community through an online consultation in early 2012.

2. National Strategy for Research Data in Canada

The first draft of the National Strategy for Research Data in Canada presented here is based on the invaluable input received at the 2011 Canadian Research Data Summit. The Strategy is composed of three parts:

- Vision: Entitled, *Profiting from the Data Deluge: A Vision for 2016*, this is a revised version of a draft vision that was distributed to participants in advance of the Summit.

- High-Level Goals: These represent the most often cited goals by participants during a comprehensive discussion at the Summit around the goals of National Strategy for Research Data in Canada.

- Framework for Action: This contains the recommended activities for the major stakeholder communities involved in managing research data and is based on Summit feedback and input from the RDS Working Group.

A detailed account of Summit input is available in Appendix 1 and 2.

2.1 Profiting from the Data Deluge: A Vision for 2016

We envision a future in which the value of our nation's investment in research is maximized and, as a consequence, Canadians and the world benefit from the tremendous social and economic advances that result. Canada maintains its position as a recognized leader in an increasingly global research environment. Researchers from all disciplines have ==*full and open access*== *to research data enabling them to conduct leading-edge research and collaborate in international endeavors.*

Canada is a country in which open data, citizen science, evidence-based policy-making, and broad public engagement with research data and science flourish. Research data are considered a public good and there is broad recognition of the value of this data beyond the research community. All sectors of society, including industry, practitioners, and the public are actively exploiting research data for commercial, health, policy, and creative purposes. To this end, research data in Canada are systematically managed, preserved, and re-used to advance innovation and Canada's leadership in the global digital economy.

Policies: Canadian organizations have coherent and cohesive policies requiring open access to publicly funded research data. The policies clearly articulate the terms of data ownership and the specific conditions under which data should not be disclosed for ethical reasons. Policies are based on agreed-upon data management principles that apply across disciplines and stages of research.

Sustainability: Governments recognize the benefits of managing research data for society and the economy, and provide appropriate leadership and funding. Research data are entrusted to an enduring institutional environment with long-term commitments to preserve and provide access to data. Collectively, the range of funding mechanisms covers the operational costs of creating, managing, and preserving research data throughout the data life cycle.

Roles and responsibilities: Roles and responsibilities are clearly defined, understood and accepted by all stakeholders in the research process. Stakeholders, through their distinct set of responsibilities, also act in partnership with other stakeholders to pursue higher-level stewardship goals important to the entire research community.

Capabilities and education: Research data management is a core component of the curricula across disciplines and all levels of education. Researchers in all sectors are well trained in how to manage their data and have access to data scientists and information professionals who have the expertise to guide management and access to research data. Working researchers have access to data management training and education resources to support their data management activities.

Infrastructure: Canada supports a national, coordinated network of repositories and services for collection, preservation and dissemination of research data, which make use of green storage facilities wherever possible. Canadian services are linked to the global research data ecosystem and are interoperable with other national, international and disciplinary networks. Data quality, integrity, and interoperability are ensured through adherence to international

discipline and metadata standards.

National coordination mechanism: Canada has a national mechanism that plays a leadership role and serves as a focal point to organize, coordinate, and support data stewardship activities across the country and bring together regional and discipline networks. This mechanism also: (1) ensures a Canadian presence in international research data initiatives; (2) houses a training and resource centre dedicated to advancing research data skills, standards and practices; and (3) provides advice on research data policies and practices.

2.2 High-Level Goals

Policies:
- Canadian organizations, including governments, funding agencies and universities, establish policies ==requiring open access to== publicly funded research data (subject to legal and regulatory requirements).
- Funding agencies include data management planning as a core aspect of research proposal review.
- Universities work with funding agencies and researchers to monitor compliance with policies and ensure that subsequent research funding is contingent on successful data management from previous ==grants.==

Capabilities and Education

- Data management training modules are implemented into research methods courses (e.g. community colleges, university graduate programs).

- Training in data management is available to researchers across Canada.

- Expertise is developed to provide support for researchers to help them with their data management efforts.

Infrastructure

- There is a common vision and clearly held understanding of the environment making up a national research infrastructure.

- A national coordinated network of trusted regional digital data repositories, data warehouses, and data libraries is established that is inclusive of all projects, disciplines, and institutions regardless of size.

- Strong international links are established and maintained to ensure Canadian infrastructure is interoperable with that of other nations.

Research Culture

- Canadian research culture recognizes the value of data sharing.

- Institutions implement evaluation criteria for faculty that consider data management performance.

- Incentives are developed that encourage publication/sharing across all stages of research, and across all disciplines and domains (government, academia and industry).

National Coordination and Leadership

- Key stakeholders come together as part of a formalized body and define a governance structure, and roles and responsibilities (coordination, policy, operational, etc.).

- Canadian priorities are identified; we build on our existing expertise, and become leaders in these areas.

- The Coordination body stewards the implementation of the Summit vision and maintains alignment of Canadian activities with relevant international bodies and initiatives.

Funding and Sustainability

- The case that research data are a national resource is well articulated through the development of a value proposition and business case for the stewardship of research data in Canada.

- A long-term funding model is developed that accurately reflects the real and ongoing costs of national research data management.

- Public-private partnership models are created where appropriate to support data management and enhance sustainability.

Standards and Interoperability

- Data management standards and best practices are broadly adopted across the country and scientific disciplines.

- Canadian researchers are encouraged to participate in the development of international standards needed to meet Canadian requirements.

- Strategies are established to integrate diverse and disparate data sets and build upon existing successful Canadian and international multidisciplinary initiatives.

2.3 Framework for Action

Modest progress is being made in several areas such as policy harmonization, development of standards, and training opportunities. In addition, Canada already has some of the pieces of a national strategy in place, including a high-speed network, national standards organizations, and selected disciplinary repositories. However, a more comprehensive approach is urgently needed.

There will have to be concerted action on behalf of many players. No one or two stakeholders can achieve the vision alone. Accordingly, a Framework for Action has been established that includes recommended tasks across the major stakeholder communities. The Framework is compromised of three goal-driven phases, with the ultimate aim of achieving the *Profiting from the Data Deluge: A Vision for 2016*.

Funding agencies, universities, governments, research communities and others all have important contributions to make. In particular, the framework recommends that:

- Canadian governments establish policies that make recipients of public research funding accountable for managing research data.

- Funding agencies require that funded projects adopt appropriate standards for data management and that data management plans are part of funding applications.

- Universities further develop the infrastructure for collecting and preserving research data.

- Researchers ensure data management standards are taken into account when creating and analyzing research data.

In addition to sector-specific activities, it was widely acknowledged at the Summit that it would be very difficult to implement the vision without some means of coordinating activities across sectors and disciplines. To that end, the framework contains a process to establish a national coordination mechanism. The mechanism, tentatively referred to as Research Data Canada, will play a leadership role and serve as a focal point to support data management activities across the country. In addition, an Advisory Council composed of high-level decision makers from a variety of sectors will be formed to provide guidance and help raise the visibility of the benefits of greater access to research data in Canada.

The Framework was developed by the RDSWG, based on input received during the Summit. It represents the collective view of the RDSWG, not the positions of the organizations that RDSWG members represent. The Framework, along with the Vision and Goals, will be further validated in early 2012, when the RDSWG will seek endorsement and further input from the broader community through an online consultation.

Table 1: Framework for Action

	Phase I: Present to 2013	**Phase II: 2013 to 2014**	**Phase III: 2014 to 2016**
Federal Government	Acknowledge that improving research data management in Canada will require coordination and resources outside of the individual research project. Establish policies that make recipients of public research funding accountable for managing research data in a manner that ensures long-term preservation and access. Support a national body that represents key stakeholders, and is responsible for executing a national strategy for research data management that addresses the issues. Review international best practices and establish a Canadian program and funding model for the creation, and ongoing support of, linked national research data repositories.	Support the design and implementation of a successful model for funding essential data management infrastructure. Establish and maintain balance between various infrastructure pieces: high performance computing, network, and data.	Establish at least one national research data repository. Establish green storage facilities for archiving of research data.
Provincial and Territorial Governments	In coordination with the federal government, establish cohesive policies that make recipients of public research funding accountable for managing research data in a manner that ensures long-term preservation and access.	In coordination with the federal government, support the design and implementation of a successful model for funding essential data management infrastructure.	Establish elementary and secondary school programs addressing data discovery and access to educate youth and generate interest in research.

	Phase I: Present to 2013	**Phase II: 2013 to 2014**	**Phase III: 2014 to 2016**
Funding Agencies	Organize expert working groups to define terms of data policies. Undertake a global survey to assess best practices in other government jurisdictions in regards to policies. Create a template for data management plans that can be incorporated into individual research proposals. Work with universities to examine the possibilities of incentives for data management, such as those for publications/citations.	Harmonize data policies across agencies. Develop guidelines to facilitate the availability of data on human subjects and other data subject to privacy, confidentiality and security restrictions. Require that funded projects adopt the appropriate standards for data management in their fields and that data management plans are part of funding applications. Undertake awareness campaigns to increase the visibility of data standards, tools and best practices (with universities).	Require that all funded research be made openly available for future use and ensure this is a condition attached to future funding decisions.
Universities and Other Research Organizations	Build awareness around data management culture on campus. Recruit local campus researchers who can speak to the data-sharing norm to help raise awareness. Work on building expertise in the area of data management on campus (i.e. through data librarians).	Begin to develop a network of digital data repositories for collecting research data. Increase the visibility of data standards, tools and best practices on campus. Catalogue data holdings and contribute to an inventory of existing data sets in Canada. Review possible incentives for research data management and sharing. Support the development of training opportunities on campus.	Maintain sustainable research data repositories. Support the implementation and enforcement of funding agency data policies. Provide support on campus for data management activities through employment of trained data scientists. Implement rewards for data management and include these in promotion and tenure processes.

	Phase I: Present to 2013	Phase II: 2013 to 2014	Phase III: 2014 to 2016
Research Communities	Ensure data management standards are taken into account when creating and analyzing research data. Participate with standards organizations and others in the development of research data management standards in their fields.	Commit to teaching students about the value of data management. Comply with funding agency and other data policies that require the preservation and access to research data.	Deposit research data into repositories and enable others to re-use that data for further research and innovation.
Summit Invitees	Endorse the revised vision and goals for a national strategy for research data in Canada and provide feedback on the Framework for Action. Support the leadership of the Research Data Strategy Working Group and the formation of a national coordination mechanism, Research Data Canada.	Participate in working groups around the various themes and contribute to the development of a Roadmap for moving forward.	
Research Data Strategy Working Group	Post the proposed National Strategy for Research Data in Canada online and seek input and endorsement from the broader community. Publish revised National Strategy and Roadmap for Research Data in Canada. Create an Advisory Council to guide the establishment of Research Data Canada. Develop a governance structure and membership model for Research Data Canada. Develop a business plan for Research Data Canada.	*[The work of the Research Data Strategy Working Group will be subsumed by the newly established Research Data Canada.]*	

	Phase I: Present to 2013	**Phase II: 2013 to 2014**	**Phase III: 2014 to 2016**
Research Data Canada (National Coordination / Support Mechanism)	Launch Research Data Canada. Form Working Groups to pursue priority activities identified through the online consultation. Develop a government relations strategy and communications plan to raise awareness of the issue. Plan a follow-up meeting to the Summit to track progress of various themes.	Organize, coordinate, and support data management activities across the country. Develop a training and resource program dedicated to advancing research data skills, standards and practices. Engage with the international community to ensure Canadian research data management activities are interoperable with other national and disciplinary initiatives. Provide research data-relevant policy advice along with other organizations.	Identify the key challenges for utilization of data across disciplines and develop a framework outlining what a multidisciplinary approach would look like. Catalogue and disseminate best practices for data management. Define a gradual route to begin merging the domain silos and bring together pilot projects. Start small, take baby steps, and begin with demonstration projects.
Advisory Council	Guide the implementation of Research Data Canada. Provide leadership and engage at a high level with policy makers and other representatives to gain support for a national strategy for research data management in Canada.	Provide guidance and advice to Research Data Canada. Raise awareness of the need for research data management in Canada.	

3. Critical Next Steps

Canada's researchers, industry, and others will become increasingly hindered in their capacity to innovate if we do not take immediate steps to implement a national approach for managing the valuable research data produced in Canada. To capitalize on the momentum of the Summit and begin to move forward decisively, the Research Data Strategy Working Group (RDSWG) will undertake the following steps in the next nine months:

January 2012 – March 2012: Consultation on the National Strategy for Research Data in Canada

The draft National Strategy presented here, which includes a Vision, High-Level Goals, and a Framework for Action, was developed with input from over 150 participants who attended the *2011 Canadian Research Data Summit*, representing a wide variety of organizations, sectors and disciplines. As such, it stands as a strong endorsement of the directions and activities documented within. However, to ensure that there is widespread consensus around the priorities and specific roles and responsibilities contained in the Strategy, the RDSWG will undertake a consultation with the broader community. In early 2012, the RDSWG will post the Strategy online and actively solicit others across the spectrum of stakeholders to contribute comments and help identify priorities. Based on this feedback, the Working Group will then publish a National Strategy that will be used to guide future activities and assist stakeholders in moving forward collectively.

January 2012 – June 2012: Establish 'Research Data Canada'

A broad consensus emerged at the Summit about the need for an entity to coordinate research data management activities across Canada, and engage at the national level with the international community. The RDSWG in its current form does not have the capacity to take on this role. To this end, the Working Group will endeavor to establish a national coordination mechanism. Tentatively referred to here as Research Data Canada, the organization will play a leadership role and serve as a focal point to support data management activities across the country.

Over the next six months, the RDSWG will develop a governance structure and membership model for Research Data Canada that will widely represent stakeholder communities. In addition, the Working Group will form an Advisory Council composed of approximately 8-12 senior leaders from government, academia, non-governmental organizations, and industry. The Advisory Council will provide guidance to Research Data Canada and reflect a high-level commitment to the issue of research data management. Once established, the first priority for Research Data Canada will be to create working groups and begin developing a detailed roadmap for moving key actions forward. As Research Data Canada ramps up it will replace the Research Data Strategy Working Group.

III. Appendices

Appendix 1: September 2011 Vision Statement and Comments

The vision statement included below was distributed in advance of the Summit. At the Summit, participants were asked to respond to two questions related to the vision: "What do you like about the vision described in the background paper?" and "What improvements would you suggest for the vision? What's missing? What must a Canadian vision on research data do for you?" The responses to these questions are documented following the vision statement.

Canadian Research Data Stewardship: A Vision for the Future

We envision a future in which the value of our nation's investment in research is maximized and, as a result, new social-economic benefits are created for Canadians. There is broad recognition of the value of research data. Researchers from all disciplines have widespread access to research data enabling them to conduct leading edge research in Canada and participate actively in international data-intensive research endeavours. Industry, practitioners, and the public are able to exploit research data, where appropriate, for commercial, policy and creative purposes.

Research data in Canada are systematically managed, preserved and utilized to advance innovation and Canada's leadership in the global digital economy.

- **Policies:** Canadian organizations have established coherent and cohesive policies based on agreed-upon national data management principles that apply across disciplines and the lifecycle stages of research. Data management plans are an integral part of funded research and there are institutional reward systems that recognize researchers' contributions in successfully producing, sharing and preserving data.

- **Sustainability:** Research data are entrusted to an enduring institutional environment with long-term commitments to preserving and providing access to such data. These institutions employ international and national standards. Collectively, the range of funding mechanisms covers operational costs throughout the data lifecycle.

- **Roles and responsibilities:** All stakeholders in the research process understand their roles in and perform their responsibilities with the design, production, management, analysis, preservation and reuse of research data. All stakeholders, through their distinct set of responsibilities, also act in partnership with other stakeholders to pursue higher-level stewardship goals collectively important to the entire research community.

- **Skills and training:** Scientists, trained data scientists and information professionals provide the necessary skills to select, manage, and provide access to research data. All other stakeholders in the research process are similarly well educated on their own roles and responsibilities with respect to data management and stewardship.

- **Infrastructure:** Canada supports a national, collaborative, interoperable network of institutional services, including data repositories, data centres, data warehouses and data libraries, that collect, preserve and disseminate valued research data, together with a high bandwidth research network that enables access. Through these services there is widespread adherence to metadata and other standards.

- **National coordination mechanism:** Canada has a national entity serving as a focal point to support data management and stewardship activities across the country. Bringing together regional and disciplinary networks; this entity (1) ensures a Canadian presence in international research data initiatives; (2) houses a training and resource centre dedicated to advancing research data skills, standards and practices; and (3) helps provide research data-relevant policy advice along with other organizations.

Q1: *What do you like about the vision described in the background paper?*

- We like that there is an attempt to describe the vision for Canadian Research Data.

- Like that it is connected to the international community and not just national

- Like it but not bold enough

- Agree on the theme topics, but there are cross cutting issues that should be addressed.

- We like the second sentence of the Vision statement regarding "…the broad societal recognition of the value of research data…"

- Ideal situation

- Great plan

- Main points have been highlighted.

- A bold vision – that outlines the need to bring multiple major players to a shared approach to data management

- Working toward defining a role for national governments in the maintenance and curation of large data sets, including the issue of data interoperability

- Draws attention to the development of critical HQP component to future data management

- It was short.

- It's comprehensive and applicable to many different areas of research.

- It has clear roles and responsibilities – the sense of co-ownership of the problem

- Reaching an understanding of distinct responsibilities is clear.

- Having a place to put the data is not enough – sustainability.

- Data is infrastructure.

- Tying research funding to a data analysis plan

- Skills and training is very important because the capacity and understanding for/of data management are weak.

- We have bled this response into 1 and 2 – tried to separate

- "Nice" vision

- Wish list – covers all the bases, comprehensive

- Does this make it "rhetoric"

- How do we get beyond this?

- Mercifully brief

- Too buzzy (buzz words)

- Whatever comes, it will fit into this

- Has something for every audience

- Is first sentence for government?

- Like the fact that it is broad, however it is a bit "magical". Could be opened up more

- Like that it is broad…

- Has the right elements, but is not a vision (the vision is already defined by the research community).

- The researchers are already setting the stage, and as a country we need to catch up

- Overall, well liked by table

- Succinct

- Recognizes that infrastructure is needed

- Comprehensive

- Concise, crisp, comprehensive

- Most of the issues are covered

- Well balanced

- It is an idealistic vision

- Pan-Canadian vision that recognizes roles

- Like first sentence

- Like the holistic approach, inclusive of all disciplines

- Addresses most of the issues

- Like that it talks about optimizing public funding

- Sustainability is key

- Vision is good. Requires a cultural shift re. Stewardship and responsibilities

- Like: Widespread access to research data

- Need for investment is highlighted in text. Benefit to Canadians is a good objective.

- Question is: how do we get there;

- Involving key players

- What's good about the vision is that it is all encompassing – if you drop anyone of the components, it will fall apart

Q2: What improvements would you suggest for the vision? What's missing? What must a Canadian vision on research data do for you?

International perspective

- For Canadians *and* the rest of the world
- Most research is international; can't have a nationalistic perspective
- Collaboration on international scale – partnerships- also coordination of effort within Canada
- Missing an international component
- The international component should be more prominent
- International role/linkages missing
- Should promote international standard rather than national
- Should be a greater emphasis on international aspect and open standards for infrastructure and sustainability
- Needs to be a global vision- the international context is an important drive
- Doesn't capture Canada's role in international community
- Is the global perspective obvious enough
- International perspective is missing- we need to work on a common way internationally for global research projects

Defining data

- Data needs to be more clearly defined
- The current plan seems to be very linear; it needs to recognize that data creation is not linear it is iterative – new data/ old data/ new analysis tools. Data is not static. Data about what data becomes i.e. data about changes in data from 24 X 7 sensors. Data is incredibly fluid.
- Question: what is research data?
- Research that has just been acquired?
- Research that has been processed?
- The results?
- In terms of infrastructure, you are dealing with different things. Results are more easily managed than the raw data. Sharing both would be of interest.
- Need to mention information as well as data, and more emphasis on metadata
- Research data definition: there are three types of research data: academic, industry and gov.

- What is the distinction between data and information
- Research data definition: raw and refined
- Research data definition: view of data is too linear-data creation is very iterative
- More explicit inclusion of all types (disciplines) of research data
- Needs a definition of research data: Data produced by research, data used in research, data derived from research

Public value

- Public value: is this really "for Canadians"?
- Where is the public in this vision
- What is role of public?
- Should focus on public access, not just access for researchers
- Need to emphasize citizen science
- What is role of political system?
- What about making it useful...accessible beyond academics
- Doesn't have anything that will sell the idea to Canadians- Canadians don't see the value in sharing data-why should Canadians care?
- Needs to promote more the value of data to Canadians
- Who is the audience for this vision? Policy makers also need to know why this is good for Canadians.
- What about the public good for data sharing, not just the economic value
- What will be the mechanisms for citizen involvement in science
- Should emphasize the public value, because research is publicly funded.

Benefits/drivers

- Canada must be leaders in research – data management is an enabler of that
- Role of government is alluded to, but not explicitly stated.
- The statement in the box when it speaks of benefits should speak to improvement in quality of life i.e. the use of data in health.
- The strategy needs to reflect not only the needs of researchers but the needs of society for research data.
- Assumption there is value add
- Economic benefit cannot be the only driver, must also talk about improvements to

"society, health, and well being".

- Other drivers include climate change, genome project
- The benefits of research data stewardship need to be emphasized
- Linkage to spatial data needs emphasis
- Evidence based decisions and government is important
- Not just socio-economic benefits, also about intellectual curiousity- what are the social benefits
- More emphasis on drivers

Ethics, confidentiality, privacy, IP
- No balance with respect to ethics, confidentiality, intellectual property. Need some wording to reflect this is required.
- Ownership is not well defined: data is not Intellectual property.
- Privacy aspects and legal issues not clear enough
- Not much discussion about implementation – perhaps a bit early to do that, just now
- Values and ethics must be clearly included.
- No mention of privacy issues
- Rights management issues need to be addressed
- Digital rights management and the role of copyright are not mentioned

Roles and responsibilities
- What is role of political system?
- Role of government is alluded to, but not explicitly stated.
- Need to clarify all stakeholders and their roles
- Actors roles and responsibilities could be better articulated
- Roles should be more explicitly stated
- "Leadership, roles and responsibilities"
- The private sector role is mission- they have money

Skills and training
- Missing focus on training young people from the start to understand the value of data, and the right ways to management data

- Instead of skills and training, use "capabilities and education"
- The vision needs to speak to the need for data science
- Researcher education not adequately covered in skills and training section, which suggests dedicated technicians, not resources

Policies

- In the policies section, data management plans should be given more prominence
- Should emphasize consistency of policies, particularly in the areas of intellectual property and commercialization versus open access

Sustainability

- Missing sustainability; hard to get grants to maintain an initiative
- Accessing and disseminating data are separate from using/analysing data
- Sustainability (big challenge), scalable, distribute

Infrastructures

- Research infrastructure is more than technologies, policies, people, and research priorities. For this to be ultimately successful, more emphasis to be placed on the people to ensure more emphasis on the "data scientists"

Resources

- Another key missing component is the resources to actually do this
- Missing element is RESOURCES

Governance

- The vision needs to reflect the need for governance of data infrastructure
- A better model would be "a coordinated network" (like the Data Liberation or Research Data Centres)

Incentives

- Rewards and recognition is missing

Data quality

I. Data quality/integrity is not mentioned

Support at the political level

- Would appreciate seeing more of a commitment in the sphere of political commitment and engaging

Specific wording recommendations

- "*Researchers… have widespread access…*" (Not sufficient), this misses interoperability and other aspects of data stewardship. Need to think strategically about all aspects of data stewardship, interoperability.

- Vision may be too narrow…should read, "…research data for the benefit of the world", don't like the "laundry list" at the end ("commercial, policy and creative purposes").

- Is stewardship the correct term?

- Would add "open and transparent" to the vision

- Concerned about the wording in the vision that highlights we need to do more, for example; enabling them to…?

- The vision needs to be clearer, the wording can be simplified, i.e.: use reused vs. optimized

- If you remove the word data and replace it with the word computing, this is the exact same wording as at Compute Canada- Doesn't seem to be anything in here that distinguishes this as unique to data

- In final sentence add, "health benefits" and policy and decision makers as a stakeholder group. Also missing are references to IPR and commercialization

- Emphasize "trans-discipline"

Other issues

- The vision statement may not adequately reflect the complexity of data management.

- Does not mention data creation

- There is an assumption that people want to share data

- Create something that doesn't create a burden – this vision should paint a world that lets me as a researcher focus on my 'priorities'

- Should focus on providing tools or support to researchers to enable the optimization of data use

- Again, this needs to be done in a more cohesive fashion so that all architectural perspectives are considered

- It needs to be recognized that this (research data strategy) is but ONE component of an overall IT architecture

- Metadata is critical

- Barriers are not explicitly stated

- Access is reflected, but the "management of data" piece, which is critical, is missing. i. e., managing data across the entire data lifecycle.

- The vision doesn't capture the full complexity of data management.

- Standardize language (e.g. Mechanism vs. entity)

- Transparency is key

- Must be inclusive of "small science" also- not just for large data projects.

- A key issue for funding agencies is data harmonization and secondary use of data

- Needs to acknowledge the differences between government, academic, industry

- Researchers want open access & easy access to data

- Openness a key concept- "full and open access" (international phrase?)

- Data quality not emphasized

- Mission statement lacks specificity

- Is "stewardship" the best word to use

- Needs a name

- National coordinating mechanism should not be part of the vision. It is more implementation that vision.

- Bottom part of vision asserts that a vision is already in place, not what should happen

- Don't like layout of the page, but like the content

- Vision statement is a bit long and will need to be shortened for politicians

- Need to include the word "innovation"

- Need to include a focus on open data

- Data generation step is missing

- Need to be explicit about connection between HPC [high performance computing], connectivity and data (3 pillars). National coordinating mechanism should include all three.

- Needs a date by which to achieve this vision

Appendix 2: Barriers, High-Level Goals, and Actions for a National Strategy for Research Data in Canada

Summit participants identified barriers, high-level goals, and specific activities for implementing a National Strategy for Research Data in Canada. Numerous ideas and recommendations were put forward, and are documented below. The input received has been organized according to 8 broad themes: policies, capabilities and education, infrastructure, research culture and incentives, coordination and leadership, funding and sustainability, and standards and interoperability, and others.

Policies

Barriers

- There is no national policy addressing data manageme
- Difficult to have a single policy. Data come in very different types. Ideal: broad policy framework, with requirement to store data. Within that, decide in each area what the realistic way to implement this
- There is no policy for data sharing in Canada. In some disciplines people share. In others they don't.
- There are privacy limitations to sharing data about human subjects.
- Who decide what is going to be published in term of data on human subjects? Ethics board?
- Existing funding agency policies emphasize intellectual property and confidentiality, not data sharing
- Need to balance privacy concerns with access. How do we include the importance of privacy, without allowing it to over-reach or become a obstacle to sharing

How can we overcome these barriers?

Canadian organizations, in particular funding agencies, must implement cohesive policies in the area of data management that require data deposit into repositories after a given embargo period. Equally important is to ensure compliance of these policies with clear rewards and consequences for non-compliance, such as no further grant funding when policies are not complied with. Researchers need clear direction in terms of how the various policies governing data (e.g. data sharing, confidentiality, and intellectual property) relate to each other and how researchers can proceed in cases when the policies seem to conflict with each other. The policies should clarify and harmonize issues around intellectual property, including limiting the time for which data are owned by the person who created them. Policies should also include the requirement that data management plans are part of the adjudication process for grant funding. In regards to government generated research data, Governments could build on the open data initiatives and implement a government-wide policy that requires that

<mark>all government-funded data be accessible</mark> (where appropriate and with adherence to standards). Policies must also be accompanied by initiatives that support adherence, such as training and infrastructure.

For your assigned theme, Policies, what are the high level goals we should set for ourselves for a national research data strategy?

- Coordinate policies across the various stakeholder organizations including funding agencies, universities and governments
- Address the competing values of open access vs. issues of confidentiality and intellectual property in the policy context
- Provide specific guidance to researchers in terms of how to comply with data sharing policies, especially in regards to data on human subjects and data subject to IP (Guidelines for data release (that address IP, confidentiality issues) (Defining types and conditions for data retention, data destruction and privacy protection)
- A recognition in the policy context that there are a wide variety of data types and data sets sizes, and that these different types may have to be treated differently-While acknowledging that big data is important, it is equally important to ensure that the small data sets receive equal attention
- Include enforcement mechanisms for policies, such as including data management as a criteria for further funding
- Policies are accompanied with infrastructure and resources so that they can be adhered to
- Include data management planning as a core aspect of research proposal review
- Ensure incentives are in place to encourage researchers to adhere with data sharing
- Recognition of contribution to data management and what kind of rights are associated with that contribution
- Policies require that data is accompanied by appropriate, open and international standards in terms of formats and metadata
- Policies require that data sets created through their funded research are registered with a specific body
- Policies adapt and evolve as technology advances and enables better protection and anonymization of data
- Policy to drive national coordination and recognizes the jurisdictional limits, drives the linkage between Canadian research communities and their individual polices and best practices, links priorities.

What actions, projects, initiatives could be undertaken in order to achieve these goals?

- Funding agencies should require that all funded research be made openly available for future use, provide the means to do this, and make sure this is a condition attached to future funding decisions

- Canada to facilitate the big journals in requiring that data associated with papers be made publicly available. They can define requirements and point to repositories

- Undertake a global survey to assess best practices in other government jurisdictions in regards to policies

- Funding agencies organize expert working groups to define data and terms of data policies (perhaps organized by funders)

- Develop privacy guidelines for releasing data

- CISTI or RDSWG work on next steps

- Commit to continue to develop federal national policies and strategies to support data management

Capabilities and education

Barriers

- Training for researchers in managing data

- Training requirements

- Repositories only are not sufficient, need expertise to manage and utilize and ensure data is accessible

- Lack of skills

- Lack of expertise in terms of using (re-using) data

- Skilled scientist in data management (HTP) – the ability to make the data understandable for the analysis; metric is usage by scientists, if there were trained data management experts (cross disciplinary is even more difficult) they would likely go unused

- Managing huge volumes of data, what to keep and what to discard – do librarians still have the right skills to do this?

How can we overcome these barriers?

In the short term, stakeholder organizations can focus on re-training of people who have some of these skills, and developing "how-to" guides, for things such as data management plans. However, in order to ensure sustainability, we must begin to include data management training as a component of university graduate education programs across disciplines. There may also be some value in introducing data management education at the high school level. In

addition, there should be an educational stream available in Canada for data science, and this should be promoted as a career path.

What high level goals should we set for ourselves for a national research data strategy?

- Design and implement a data management 101 course and ensure that it includes a component

- Participate in promoting data management skills in the international environment

- Work to implement data management modules into research methods courses at universities

- Training – consistent training for Canadian researchers, to drive standardization. National training standards for data managers; multi-tiered (e.g. community colleges, university grad programmes)

- Provide expertise for research to help them with their data management program

What specific actions, projects, initiatives could be undertaken in order to achieve these goals?

- Enhance the CARL course on data management and expand participants to include researchers (not just librarians and data specialists)

- Recruit 5 or so people to develop an implementation strategy for training

- Move geospatial data management forward and continue to interact at the national and international level (UN, GEOS)

- In the short term, facilitate knowledge exchange and aggregate learning until more formal training standards are available

- Institutions run pilot projects in specific disciplines with little track record of sharing and training programs

- CERCs: recruit a data expert. Tag one of the chairs for data management and data storage, the same way a chair was dedicated to Automotive research

- Skills and training initiatives: develop a national training program

- Steering group representing CAGS, tri-council, AUCC to discuss a national training strategy for research data management training at the post graduate level (perhaps add these reps to RDSWG).

Infrastructure

Barriers

- Lack of Infrastructure

- Tools and Infrastructure

- There is not what we need for our volume: capacity, bandwidth;

- Lack of repositories

- Infrastructure

Goals

- Cohesive network of research repositories

- Create a national infrastructure

- National repository for research data – distributed environment

- A clearly held understanding of the environment making up a national research infrastructure

- Inclusive involvement of the infrastructure; available to all projects and institutions regardless of size

- Green storage supported by CANARIE allowing researchers to capitalize on available storage

- Infrastructure to support open data and data preservation

For your assigned theme, Infrastructure, identify two or three high level goals we should set for ourselves for a national Research data strategy

- Study and develop methodologies and international standards on data disclosure, with other countries

- International best practices on academic integrity

- Inventory of what data is available in Canada right now (how many silos do we have?)

- Need a project manager and champions to drive the project

- Buy-in by politicians

- CARL has an application going to CFI and so do others. We need a common voice to support one another, and endorse one another and complement one another.

- Increase the visibility of data management issues

- A forum to facilitate exchanges, etc.

- Hold data conferences with an international flavour

- Meet again in 18 months

- Projects focused on specific segments of the data management issue (big picture, with smaller working groups)

- Need leadership

- Form working groups and report back next year

- Build the community now to support each other. We are all struggling in the current political economic environment

- Establish a national coordinated system of trusted digital repositories

- Establish a workbench for finding, sharing and using data- archival function, data conversion tools

- Create a Canadian Research Data Curator/Facilitator/Data jockey (similar role to a Canadian Research Chair)- to champion and play a data stewardship role

- Stable funding for infrastructure

- Include domain experts as part of the infrastructure

- Clarity of roles amongst the various stakeholders

- Coordinated infrastructure, but distributed across Canada, and division of labour according to discipline- and interoperable internationally

- The infrastructure could possibly have basic and then other services (that could be charged for)

- Stable funding commitment

- Integrate the infrastructure pieces (data, computing, network) into a single IT infrastructure entity and linked internationally

What actions, projects, initiatives could be undertaken in order to achieve these goals? Who is best positioned to lead? What tools, incentives and approaches are necessary?

- Submit CARL CFI application for data preservation repository network

- Private sector and how public sector will engage

- VPs Research building issues and awareness around data management culture

- Tri-councils set the data management context through policy

- Find local campus researchers who can speak to data sharing norm

- Facilitate a network of organizations that create, update, and access data

- Define methodology and standards on disclosure of clinical trial data

- Support the CARL CFI proposal for data repository infrastructure- get commitment from the universities for the proposal- make sure VPs Research know about the proposal and are supportive of it- draft a letter of endorsement

- Use conference email list to send out common messages to the VPs Research

- Get in touch with Data Librarians at the institutions and work on building expertise in the area of data management

- Expand role of data librarians

- Create metadata repositories to bring metadata up to certain standards

- Awareness campaigns to increase the awareness of data, tools and best practices. Both funding agencies and data librarians have a role to play.

- Endorse a vision

- Propose a roadmap

- Develop an inventory of data repositories (discipline, project, and government-based)

- Identify 10-20 nodes (perhaps disciplinary)

- Granting councils require data management plans as part of funding applications, perhaps also create a template for these

- Granting councils develop incentives and enforcement mechanisms

- Prioritize areas of research that aren't well covered

- Work with provosts, faculty associations and scholarly societies (and CAUT) to consider how to promote the culture of data sharing

- Rewards for data management included in promotion and tenure processes- perhaps developing tools using DataCite

- Collection development policies for data because we may not be able to collect all data

- Principle: If you receive public funds, then you have an obligation to share (with or without an embargo period)

- University libraries are working with CANARIE and Compute Canada to develop a system of repositories, with possible funding from CFI

- Environmental scan and best practices

- Granting councils create and fund a working group that have a mandate to form an action plan and strategy

- Each stakeholder contributes to the RDSWG and it becomes a membership organization

- Annual conference at which participants report on their progress

- Communication plan to raise awareness of the issue

- Build research data repositories in support of data management activities

- Membership on committees and active on WGs

- In kind staff contributions from all stakeholders

Research culture and incentives

Barriers

- The current academic promotional and tenure system does not support sharing data
- There are no academic rewards or other incentives for data management or sharing
- There are differences in IP rights across disciplines
- Researchers and others who collect data think that they have exclusive rights to it
- We have no mechanisms for mediating ownership rights
- Competing values of protecting IP and ownership vs. providing open access to publicly funded data
- Researchers don't want to share their data in fear of being scooped
- There is an emphasis in the research environment on IP, not data sharing
- There is little willingness to share by both governments and researchers
- Institutions are reluctant to put forward incentives

How can we overcome these barriers?

In addition to policies, researchers need incentives to share their data. Universities and other research institutions should implement systems that reward data management and data sharing. Reward systems differ according to discipline; therefore different types of incentives may need to be implemented in different fields. However, data management activities should become part of the promotion and tenure criteria, and included in other career assessment milestones. A generally accepted citation scheme should be developed, whereby the original creators of data are acknowledged by others who use that data.

High-level goals

- Change the research culture from one of competition to cooperation/collaboration and to recognize the value of data sharing
- Link to training for researchers in managing data
- Governments address the question of privacy and personal information in order to develop a Canadian policy on IP that is "ready for the future"
- Acknowledge that data management is a part of Indirect costs
- Institutions change evaluation criteria for faculty to include data management activities
- Reward data deposit into repositories
- Develop a set of incentives that would encourage publication/sharing of data. Must act all across the board, i.e. at all stage of the research and across all disciplines and

domains (government, academia and industry)

Leadership and coordination

Barriers

- The landscape is very complex, for example provincial jurisdictions vs. federal. How can deal with this complexity?

- Communication between the various research communities, institutions, agencies, department

- Provincial, institutional, cultural, individual, political boundaries and barriers

- Canada does not act like a team (e.g. CFI only pays 40%, provinces pay 40% ONLY if infrastructure is in the province- whereas the infrastructure should be located in an area that is most conducive to the issue at hand

- Lack of cohesive leadership makes it difficult to build consensus on best practices – thus different groups are arguing different approaches and there is no-one focal point to resolve these issues

- How to coordinate multi-disciplinary activity

- Lack of coordination and leadership

- Linking and coordination with international community

- Fragmentation of effort for data management

- Where could sustainable funding come from? Pitch directly to federal government?

- Lack of a framework, for a researcher to fit into, that guides and supports in a way that enables the principles / aspects of the vision

How can we overcome these barriers?

- The federal government has an important role to play in coordinating all of the stakeholders involved and in moving the vision forward. This should include designing a governance model that allows decisions to be made and directions to be taken that are reflected across all Canadian by all research entities; and ensure that there is appropriate support for things such as the development of data management standards and infrastructure architecture. The universities are well placed to take on responsibility for data preservation over the long-term. And, there is a role for industry in terms of development of tools for data storage, processing and reuse. Although there is not yet consensus about what type of model would be best for Canada (centralized vs. distributed; cathedral vs. bazaar) there is still a need for some central coordination. Must mesh and be interoperable with other national strategies and approaches

For your assigned theme, National Coordination Mechanism, what are the high level goals we should set for ourselves for a national research data strategy?

- Adopt a national coordinated approach which reaches across different domains (i.e.: academia vs. government vs. Private)

- Develop an integrating strategy between diverse and disparate data sets and their associated disciplined; perhaps by modelling upon existing initiatives that are inherently multi-disciplinary and that have achieved some of success in integrative informatics ventures.

- Bring key stakeholders together and set a governance structure

- Identify national leadership and champions to ensure it stays on track and is prestigious

- Implement the National Coordination Mechanism with the 3 roles as outlined in the background document: (1) ensure a Canadian presence in international research data initiatives; (2) house a training and resource centre dedicated to advancing research data skills, standards and practices; and (3) help provide research data-relevant policy advice along with other organizations

- Identify Canadian priorities, build on our existing expertise, become leaders in these areas, and share with other international initiatives

- Perhaps engage with outside organizations like Google.

- Create an entity to steward the process. Kind of "Research Data Canada". A body to implementing the vision expressed in the document.

- Collaborate across disciplines and sectors within Canada, and also develop partnerships on an international scale

- Develop a framework outlining what a multidisciplinary approach would look like

- Establish and maintain balance between high performance computing, network, and data (breakdown silos)-national forum to bring these communities together (e.g. UK data forum or DCC)

- Develop a solution that is inclusive, comprehensive and is accompanied by an appropriate funding model

- Define a gradual route to begin merging the domain silos- bring together pilot projects

- Put in place something – some "data management" body/strategy - that stays despite rotation in government

What actions, projects, initiatives could be undertaken in order to achieve these goals?

- Acknowledgement that we need new models for tackling these problems. Look to

other jurisdictions (i.e. Australia) to learn from their experience

- Design a model for success for sustainability

- Evaluate RFP approaches to moving forward

- Create an umbrella coordination mechanism to guide post summit activities. In partnership with key stakeholders (funding agencies, governments, and academia), establish a task group to define the requirements for an RFP to establish a national coordination mechanism for research data management

- Create a steering group to carry forward several of the steps within today's vision, incorporate the input into the vision received today and move fwd with that updated vision.

- Expand the Research Data Strategy Working Group (RDSWG) to include representatives from research; this group should continue to move each theme forward.

- RDSWG needs to assign tasks to organizations who have skills in particular areas – need to understand what groups exist and what their skills/experiences are, look at international research networks, how have they created/removed barriers

- Develop a government relations strategy; someone from the RDSWG (perhaps the chair) starts building relationships with government to get this on the national agenda.

Standards and Interoperability

Barriers

- In some fields, there are no standards for data management and sharing

- There are varying degrees of adherence to standards, where they exist

- Metadata and other

- Standardization – different levels, for example; Canadian standards, vs. international standards

Overcoming barriers

- Adoption of standards to assist with interoperability

- Interoperability across disciplines

- Interoperability and standards

High-level goals

- Adopt standards and best practices where solutions already exist.

- When developing data collection software, ensure data management standards are

taken into account.

- Identify the key challenges for utilization of data across disciplines
- Different institutions and bodies adopt common standards for metadata to facilitate interoperability.
- Provide access to 'data dictionaries', best practices and a forum for researchers to go for guidance about data standards.
- Develop best practices for interoperability in each discipline.
- Global Adoption of the processes, standards, culture changes by the research community.
- Define mechanisms needed to support integration, mining and access to data- in a cost effective way.
- Develop a semantic search capability across data sets.

Others: Data Ownership /Intellectual Property

Barriers

- Inertia and traditional view of data. People who collect data think that they have exclusive right.
- Issues of ownership – no way to mediate ownership rights – trust
- Changing the culture – individual, institutional, funding agencies, government
- One of the cultural barriers includes the desire to be first in research results
- Protection on IP and ownership of data vs. open access to publicly funded data
- Researchers don't want to share their data in fear of being scooped
- Ego and ownership
- There needs to be a cultural shift in the research community- there is an emphasis on IP not data sharing

Others: Privacy and confidentiality

Barriers

- There are privacy limitations. Legal issues.
- Who decide what is going to be published in term of Ethics board?
- Legal barriers, confidentiality
- Balancing accessibility with privacy
- Security and privacy issues

- Existing funding agency policies emphasize intellectual property and confidentiality, not data sharing

- Address privacy issues

- Question – how do we include the important of privacy, without allowing it to over-reach or become a obstacle to sharing

- Need to balance privacy concerns with access

Other: Diverse

High-level goals

- Social benefit: good social outcome

- Evidence based decision making

- Reusability

- Better research is important to mention.

- (Government people should be represented on the internet management forums on data.)

- Distinction between goals and vision: goals pertain to implementation; vision is an end state

- Should advance research

- Ensure data is usable in the future

- That the strategy be endorsed by the major stakeholders influenced by it

- That the strategy lead to tangible action

- To be able to increase access to the data

- Public uptake/interest in the vision. How to make the discoveries that are made through research resonate with the public and get them to realize how research and data management support the technology they take for granted

- Get the public excited about the provision to and access to research data.

- Canada must be leaders in research – data management is an enabler of that

- Needs to be a national priority

- High level goals versus making it happen

- Canadian research data become well managed and accessible

- Evidence-based decision making and policy

- Collaboration and competitiveness must coexist

- Ensure our vision is all encompassing - look to leaders in sectors where there is

experience

- All data collected through public funds be made available
- Create interdisciplinary teams to focus on specific problems
- Continuous dialogue- hold conferences (add international flavour)
- Private sector involvement in developing data management solutions
- Implement a "national centre for research curation"
- Scientists can access and use data sets easily
- Researchers can focus on research, but data is preserved
- By having a research data infrastructure, research is more efficient, scientists are doing science, and data managers are managing the data
- Ownership, and stewardship – resolving this issue should be a goal of the strategy
- Information Governance – good catch phrase
- Learn from other national initiatives (e.g. Australian initiative is government funded but run by the universities)
- Needs a new model
- Greater awareness of the urgency of research data stewardship- what will be lost if we don't
- A consensus that all research data are a public good and a guiding principle
- Ensure that IP is restricted through terms and conditions for restricted access

Other: Diverse

What actions, projects, initiatives could be undertaken in order to achieve these goals?

Making the case

We must develop a business case for data management that includes proof of value for public and for the research community. The business case will include opportunity costs of not doing anything and answer the "so-what" question.

Business case/value proposition

- To be able to define the value proposition for being to create, maintain and distribute data across all domains
- Improving the perception of a return on investment from publicly funded research
- Articulate the Return on Investment- including how much invested and what are the lost opportunity costs

- Value proposition for individual researchers

- Tell engaging stories to demonstrate the value proposition

- We need to put together a value-proposal. What is the value of sharing? What is the value of storing? Need to transform Canadian successes into a discourse that demonstrate the value of data handling.

- Use a proper data management strategy to leverage development of new economic area: development of algorithms, ontology, search engines, etc.

- Tie the short and long term Canadian financial well being to the efficient use of research data, be it because of re-use, re-purposing, re-leveraging, data sharing, etc.

- Value proposition: What is the business case for government departments to share data+

Demonstrator projects

- We should begin with demonstrator projects where Canada is already well positioned, where the communities are already engaged in this issue, and there is a high likelihood of success, for example data resulting from research in the arctic or astronomy research. Demonstrating progress gradually will build support from bottom-up and build the trust required for others to hand over their data.

Build on existing projects

- build on what is already underway; needs to be sustainable; operational as well as infrastructure; three years (or other target date to be determined) from now enumerate all data produced from tri-council funding; need quantitative information for design of solution and for measuring success; building on strengths; training; aggressive timelines w concrete deliverables

- Something concrete and implementable

- Look for models and build on what exists (e.g. IPY project)

Learning from others

- We can learn from what other more advanced jurisdictions are doing and show how other countries have benefited financially from implementing the infrastructure.

Raise awareness of importance of research data management with government

- Is the office of the minister sensitive enough on the issues of data management?
- GOAL: sensitize political leaders to the importance. IC has an interest. But what about

the other ministries? This is a pan-government issue.
- Sensitize decision makers about the need for a national and stable strategy;
- Put in place a process including an awareness campaign to ensure data sharing and release policies are drawn up and shared before implementation