THE LEADERSHIP COUNCIL
LE CONSEIL DU LEADERSHIP

Canadian DI Environmental Scan:

A Supplement to the Background Précis Document Provided to DI Summit 2012[1]

Contents

[1] Prepared for DI Summit 2014 by the Project consulting team.

Executive Summary

The following analysis, presented as both a statement of existing problems and a "SWOT" assessment of the Canadian DI ecosystem, reflects broad input from the community ,albeit not necessarily a full consensus on the framing of issues. Nevertheless there is a very significant degree of agreement on the major deficits in the research environment at present and the key areas for change. This is coupled with a remarkably strong understanding of the importance and power of collaboration, as well as the necessity of individual/organizational commitment to changing how we plan, deliver and manage Canadian DI.

Acknowledgements

This assessment of Canada's digital infrastructure (DI) strengths, weaknesses, opportunities and threats (SWOT) owes much to the excellent work that has been done over the last few years by Kathleen Shearer and the Research Data Working Group (now succeeded by RDC), CARL, those who participated in the 2011 Data Summit and in DI Summit 2012, the many participants in the annual CANARIE User Forum meetings, and the earlier national discussions on data archives, c3.ca, and CANARIE.

The Constituent Parts

Canada has many of the elements required in the DI ecosystem. Key players in the provision of DI are the funding agencies, service providers and institutions.

The federal funding agencies play a large role in the funding of digital infrastructure, with an increasing involvement of the provinces. The division of responsibility has evolved since the formation of CFI in 1997 and Genome Canada in 2000. Over the past 15 years, CFI has gradually taken on the role of prime funder for computational resources that are above and beyond those accessible from a project grant. Initially this was confined to capital and early operating costs. Most recently through CFI's Major Science Initiatives (MSI) program (currently time limited) it has taken on responsibility for ongoing, but matched, operating support of Compute Canada.

The major service providers in Canada are CANARIE, the regional ORANS, and Compute Canada. CANARIE and the ORANS are recognized as an essential service in their provision of a broadband ultra- high-speed network, with local connections and associated services. CANARIE has recently taken over management of the Canadian Access Federation - a trusted access management environment (unique-identity access) for Canadian research and higher education communities. Compute Canada delivers those aspects of advanced computing that are more effectively and efficiently delivered through a nationally coordinated body and that complement the facilities and services provided by allied organizations. This includes a portfolio of expertise, services, customized software solutions and high performance computational capacity. As an organization, Compute Canada is going through a major restructuring phase of its governance and operations.

Three other specialized service providers are i) the Canadian Research Knowledge Network (CRKN) that deals with site licenses for primary literature (and which may evolve in the future to additional digital products); ii) the National Research Council of Canada that manages DataCite, Canada's data registration service – the provision of a mechanism for registering research data and assigning digital object identifiers (DOIs) to them consistent with international conventions; and iii) CASRAI, a standards organization that deals with research administration data that is expected to include standards for metadata.

There is no single body responsible for infrastructure and tools to manage research data. However, increasingly, universities recognize that digital content is a valuable asset requiring appropriate attention.

Libraries have had a long tradition of supporting and providing access infrastructure for social science data (through Nesstar servers), but are now becoming more involved in issues of curation and preservation. Other national domain-specific specialized research data management infrastructures (RDMI) exist, including i) the Canadian Research Data Centre Network (CRDCN) that provides services, some curation and protected access to Statistics Canada data and is now expanding to include other data sets; ii) the Canadian Polar Data Network (CPDN) that is an outgrowth of the very successful research data management activity that was developed to support the International Polar Year (IPY); iii) the Canadian Astronomical Data Centre (CADC) that is an internationally connected and trusted repository for astronomy data collected through projects involving Canadian researchers. Similarly, project level RDMI exist, such as Ocean Networks Canada , but their data management and production of data products are made enormously complex and more costly by the diversity of non-interoperable standards across their research community.

There is currently significant ongoing activity in Canada relating to researcher needs and the evolution of the DI ecosystem – from major national initiatives to smaller institution and domain specific initiatives. An attempt to map some of these is provided in Appendix 1, along with a description of the more targeted initiatives. This is an indicative, but incomplete, inventory.

Appendix 2 provides an overview of the various players (other than the funding agencies) that has been extracted from the recent consultation document of TC3+ (the Granting Councils and CFI).

Framing the Problem

In 2012, a senior member of the academic community described the Canadian digital infrastructure ecosystem as having:

- Fragmented approaches
- Overlapping jurisdictions

- Multiple voices
- Inconsistent funding
- Focus on equipment rather than people
- Little attention to data as “infrastructure”
- Policy gaps

This was a provocative commentary that reflected the frustration of many who had debated the lack of our readiness for data intensive research over many years. Notable is the fact that these comments focused more on the lack of coherence and cooperative action, policy gaps, and inadequate attention to critical elements, than on funding levels. Financing of DI is necessary, but definitely not sufficient, and is not currently the major roadblock.

The work of the Leadership Council addresses the substantive issues of cooperation and balance of attention; it is first and foremost concerned with tackling the most time-critical problems identified in the Roadmap:

Governance/coordination
There is a pressing need for greater collaboration between the funders of research (the TC3+, Industry Canada and the provinces), the providers of national digital infrastructure, and the critical partners in Canada’s research enterprise, including universities, professional associations, researchers, libraries, national standards organizations and other research performing and user organizations. The collective actions of the TC3+ are an important step in this direction; but without increased collaboration and coordination, we risk fragmented approaches, sub-optimal alignment of activities and investments, and serious gaps in the digital infrastructure. While structural change is not necessarily the answer, there have been extensive and strident pleas for a more structured approach to dialogue, coordination and integration.

Policy and planning framework
Canada lacks a cohesive national policy that provides an integrated planning and funding framework for all the elements of the DI ecosystem. The lack of policy breeds a lack of strategy and the concomitant system problems. As an example, two important components of Canada’s digital infrastructure, Compute Canada and CANARIE, are funded under different structures and for differing time periods. There are asymmetries in mandate and performance expectations. There is a lack of clarity as to what is needed and how infrastructure for research data management should be effectively delivered and aligned. The lack of appropriate planning and investment horizons for all elements of the DI ecosystem engenders a “short-termism” of approach with investment cycles that are out of step with the pace of change in research and technology development, and with many DI elements developed through small-scale independent projects rather than at a national infrastructure level.

This presents a challenge in how we ensure that these foundational infrastructures deliver maximum benefit at any given time. Investments in the DI ecosystem have too frequently been project-based, creating disincentives to long-term system planning. Incentive structures have not fostered pan-Canadian collaborations.

Data management
Research data management may be the weakest link in the Canadian DI landscape, despite the massive increases in the amount of data being created daily through the research process. There is currently no policy framework, nor an agreed-upon strategy and/or the capacity to protect this valuable public asset primarily funded through public monies. Equally, there is little capacity to support access, use and reuse by a wide range of users. To date investments in digital infrastructure have been more focused on technology, without concomitant attention to data, the provision of skilled personnel, and relevant software development. The report from the Summit 2011 and the work of the RDC has laid a serious and comprehensive foundation for this; the recent discussion paper from the TC3+ is a significant move in addressing some aspects of this problem.

A SWOT Analysis of the Canadian DI Ecosystem

Underlying this articulation of problems lies considerable work over several years that has identified strengths, weaknesses, threats and opportunities for the Canadian DI ecosystem. The present document does not pretend to reproduce the extensive analysis of issues in earlier documents, but rather to distill in a few pages the current state of the Canadian DI environment, drawing heavily on those sources, and positioning that analysis in the research environment of January 2014. It is presented in the form of a SWOT analysis that extends the above problem diagnosis.

Strengths

A culture of collaboration – There is a very real culture of collaboration and recognition of the strength in working together rather than indulging in fragmentation of efforts, even with the number of diverse stakeholders in the Canadian research enterprise. Organizational players are seriously disposed to collaborate and have worked to avoid duplication, even in the absence of a policy or funding framework that makes this happen.

The leadership of the TC3+ – The actions by the TC3+ to advance policies that will inculcate a culture of data stewardship represent a major step forward in creating the conditions for optimizing the value of data produced through publicly funded research in Canada.

The *existence of the Leadership Council* – This body represents the first semi-formal Canadian attempt to explicitly coordinate the needs and activities of the many diverse bodies – across all disciplines and sectors – required to form an effective Canadian DI. It can now provide the backbone of DI design and implementation following guidance from participants in Summit 2014.

Two critical infrastructures – Canada is endowed with two key infrastructure providers that are unique and effective collaborations of federal and provincial/regional players – CANARIE/the ORANS and Compute Canada. They provide essential network and computational services in a digitally-intensive research environment. Increasingly they are working together to ensure effective interplay of their services and the provision of the middleware necessary for efficient access and use by the research and private sector communities.

Designated support for research infrastructure – The creation of CFI with its designated support of research infrastructure is widely recognized as having been transformative in its impact on the academic research environment.

Strength in data analytics – Canada has a competitive advantage in data analytics. Companies such as Oracle and IBM consider that there are significant Canadian computing and analytic strengths in both the private sector and academe.

Weaknesses

Lack of governmental engagement in policy development – There has been an absence of governmental engagement (federal and provincial) in policy relating to a national digital infrastructure. Federal-provincial relations confound this picture.

An imbalance in infrastructure elements – While there has been a collective Canadian investment in the network and computational "legs" of the DI stool, there has yet to be sufficient attention to research data management (RDM)[2] and the associated human infrastructure. In both RDM and utilization of advanced computing, the largest challenges in the DI are, in fact, human.

Funding structures - While CFI has been enormously effective at funding competition-driven support for research infrastructure (and this must, of course, continue), the funding model that pits project specific infrastructure against research infrastructure for system-level support (e.g. research platforms and generic research resources) has compromised the development of an integrated and sustainable DI. It is unstainable as a model for support of Compute Canada.

Funding levels – Finances will always be limited; as such priorities need to be set. However, the lack of an adequate planning horizon for all aspects of digital infrastructure has resulted in less than optimal allocation decisions; e.g. : i) between capital investment and maintenance costs; ii) among technology, software, and human infrastructure; and iii) among the three main pillars of DI – network, computation and data management.

The research culture – The Canadian research culture focuses more on protection and utilization of self-produced data, than on stewardship–i.e., promoting the sharing and reuse of data. There are tensions between sharing and openness, with relatively few mechanisms for mediating ownership rights in the modern data-sharing environment. To date, it has been

[2] Recognizing the four phases of RDM – production, curation, long-term data management, discovery, and repurposing/reuse.

publications, not curated, accessible and shared data that are the recognized hallmarks of academic productivity, and few or no explicit rewards for data integration and sharing.

Skills and training – There is a significant unmet need for skills upgrading, training and mentoring in the use of advanced computing, especially in disciplines that have not had extensive engagement in data-intensive research until recently.

While improving, there is still a paucity of awareness of RDM principles and good practices among researchers and research institutions; relatively few researchers have training in RDM; there are few positions for data managers/professionals; training opportunities are sparse.

Data repositories –With the exception of some discipline areas, there are still relatively few long-term data repositories in Canada (i.e., sites that manage all critical functions related to data curation, preservation and access).

Regarding non-discipline-specific repositories, OCUL is moving towards trusted data repository status with broad acceptance of research data. Further, some institutions are developing their own repositories, but they are often "dark" (i.e. not accessible to researchers outside the institution), suffer from insufficient trained personnel, and are not yet linked into a national network.

The lack of a coordinated preservation infrastructure is an obvious and embarrassing gap for a country with the research record of Canada. There is an emerging consensus that much research data management in Canada will be distributed in nature through local service sites. However, this does not obviate the need for some form of national data repository, such as: i) a redundant site for major data collections where loss of the primary site could be a serious risk, and ii) a long-term repository for data of national interest beyond the career of a researcher or an institution's commitment to hosting. As yet there is no such repository, nor methods devised to optimize efficiency and effectiveness among national, regional, disciplinary, and local data repositories.

Software and Middleware – There has been inadequate attention to the development of the software and middleware needed to make infrastructure work according to users' expectations, and usable across multiple public and private resources and platforms. Even with the new CANARIE programs, a better system-wide approach for the development, adoption, and support of middleware is required, including a mechanism to evolve integrated development platforms. Similarly there are inadequate mechanisms for innovative software development – leading researchers to rely on in-house efforts of varying quality and effectiveness (to say nothing of redundancy) or software developed in other jurisdictions.

Experimentation and research on RDM – While there is an increasing number of project- and/or regional-specific initiatives and collaborative institutional initiatives that entail research and development, there are relatively few incentives and rewards for such initiatives, especially as they are normally a non-traditional academic activity. Yet, tools that enhance the potential

usability of data have the potential to add significant value to research data that has already been funded by the public purse.

*Research **on** infrastructure* – There is a serious deficit in Canada of research on infrastructure - i) research focused on the development of innovative software and middleware that would both support research needs and support the Canadian ICT industry in developing innovative products (compare with the NSF targeted programs); and ii) reflective enquiry on the history and implications of how Canada has funded DI infrastructure (a historical and science policy approach) and what are the implications, to foster learning from history, and even more – international experiences.

Research data standards – There is variable adherence to international standards throughout the data lifecycle, with a few exceptions such as the social sciences (that use the DDI metadata standard) and "big science" (e.g. astronomy and particle physics). Discipline-agnostic standards that are of particular importance in interdisciplinary research do not yet exist.

Policy deficits – There is a paucity of policies on the full life cycle of RDM at the level of institutions and a complex policy environment at the level of both funders and institutions. Conflicts or perceived conflicts among policies currently preclude effective data sharing, even with appropriate application of privacy and confidentiality requirements.

Roles and responsibilities – There is a lack of definition and acceptance of roles and responsibilities for the DI ecosystem, in particular (but not uniquely) for long-term custodial responsibilities. Similar situations exist for provision of computational capacity among institutions, regions, researchers and Compute Canada (CC), as well as in respect of direct applications to CFI vs provision of services by CC.

Institutional (and some regional) networks – There is insufficient volume, capacity and bandwidth, and connectivity (the last mile), especially for data intensive fields of research such as genomics.

Opportunities

Underpinning for Canada's innovation environment – A robust integrated digital environment will support a strong digital economy and position Canada as a nation successfully leveraging its digital advantage.

Positioning research data as a national research resource – through improved stewardship of research data, those data produced through publicly-funded research will become more easily and routinely available for broader use, reuse and validation. The ultimate objective of good data management is to enhance its potential for further research, beyond the project through which it was generated. By improving management of publicly funded research data through the full lifecycle of asset assessment, risk management and preservation, those data will be available for use by the larger community and represent new research, innovation, and education opportunities and cost efficiencies. Such stewardship is also part of the responsible

conduct of research, ensuring that data in support of scientific claims remain available for review and challenge by the community up to a certain time after publication.

Interoperability – Interoperability (and thus cost efficiency) will be enhanced through common standards for middleware across disciplines. Such standards must be aligned with international efforts to ensure that reach is both national and international.

The Research Data Alliance – There is an opportunity for Canada to formally join the RDA, an international collaboration on research data standards that offers a rich opportunity to contribute to and benefit from the work of other nations, as well as inform the implementation of research data standards in Canada. Research Data Canada is an organisational member of RDA and could become the vehicle for Canadian full participation if the TC3+ and/or Industry Canada should decide that Canada should become a supporting country. In the three founding regions/countries, the United States, Australia, and the European Union, there are national RDA's through which the country's efforts to support RDA internationally are focused. In both Europe and the United States, these national organizations were expressly created for the purpose of national engagement in the global organisation.

Policies, incentives, rewards and recognition –The scene has been set for implementation of more universal approaches to good data management under TC3+ leadership. The current TC3+ consultation document proposes firm policies on data management and sharing, along with integration of RDM plans as a component of peer review. Federal funding leadership may well trigger broader changes in institutional reward systems.

A strong federal policy for DI would also frame the development of more effective strategies for the planning, funding and coordination of national DI platforms.

Closer alliance of Compute Canada and CANARIE – Recent discussions between Compute Canada and CANARIE, and the preparation of complementary but linked strategic plans, offer the prospect of a much tighter working relationship and enhanced coordination of service delivery to the research community.

The CARL RDM Network Initiative – CARL, together with CRKN, is facilitating the development of a national network of library-based research data management services. While still in an embryonic stage, the initiative has broad support among the library community and is being designed to offer, potentially, three functions: i) education and training; ii) services such as access to expert advice, data visualization and modelling, and data rescue; and iii) tools and technology for dissemination, discovery, preservation, and access, perhaps in collaboration with Compute Canada. Realization of this opportunity will require funding in addition to what has already been committed by CARL.

The Western Preservation Consortium – Following work with a local company Artefactual Systems Inc., UBC, the University of Alberta, and SFU have begun collaborative work on an RDMI and the establishment of a regional preservation backbone in Western Canada. As they identify and implement global standards and best practices, the expectation is that this work

could lead to a Canadian digital preservation network. This infrastructure has the potential to complement the services network being facilitated by CARL.

Research initiatives – There is a significant opportunity for Canadian benefit in modest investments in research on infrastructure – i) fostering research on new techniques and technologies, such as research on software/middleware and on managing, analyzing, visualizing and extracting useful information from large, diverse, distributed and heterogeneous data sets (as NSF has done) and ii) understanding the implications of how Canada has and is supporting DI – through commissioned and sponsored research and workshops (again – NSF serves as an example - see the 2007 workshop "History and Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures" workshop http://deepblue.lib.umich.edu/handle/2027.42/49353) .

Brokering access to commercial services – There is an opportunity for more cost-effective provision of DI services through effective utilization of commercial service provision.

Enhanced inter-sectoral collaborations – Researchers in the private sector and government have comparable needs for access to high speed networking and advanced computational capacity, as well as effective tools for research data management. There are as yet untapped opportunities for sharing DI and the concomitant benefits of enhanced interactions among academic, private sector and government researchers.

Integration of Northern interests – A stronger and more integrated DI should look to how services to, and engagement of, communities and researchers in the North can be fostered. Both the network and the collaborative opportunities are potential tools for Arctic sovereignty.

Threats

Complacency – There has been intense frustration among many stakeholders borne of much the lack of action on the problems articulated above. The lack of definitive action by organizational and institutional leaders and funders will have a major detrimental effect on the overall research environment. As one researcher noted, we urgently need to "create critical mass out of critical mess."

Lack of recognition of shared responsibility – No one player alone can be expected to deliver a coherent and effective DI for Canada, nor can it happen without clearly defined roles and responsibilities for DI coordination. The threat is of a "pass the buck" allocation of responsibility to others, rather than serious commitment to negotiating a realistic division of responsibility and financing. Quoting again from the consultations in preparing for DI Summit 2014, "United we grow, divided we status quo".

No designated locus of convergence – As St. Arnaud and Therien said in their 2012 report, *"It is hard to see how the pieces of the puzzle can be made to fit together, locally and globally, without having an explicit and well-recognized locus of convergence where these pieces can interact."*

Lack of leadership - In the absence of leadership and commitment, from institutions, governments and the Coalition of stakeholders, there will be little prospect of change from the status quo.

Appendix 1 – Map of Current and Ongoing Canadian DI Activities

The chart on the following pages maps current DI-related activities that are underway in Canada over the time frame of late 2013 to the end of 2015. Many of these initiatives would benefit from a forum for coordination and information exchange, including best practices. Interestingly, this is very much a moving target!! Note – this is an indicative list of initiatives that exist and are underway in Canada; it would benefit from a more fulsome input to make it fully current.

Following the chart is a description of those initiatives that are not captured in the descriptions of the stakeholders in Appendix 2 and that are more targeted in nature.

Charting Canadian Initiatives Relating to Digital Infrastructure

This is a current picture of the various initiatives underway in Canada with linkages to the work of the Leadership Council

Time lines reflect information available as of January 2014; subject to adjustments/refinements

X = activity; D = decision; P = policy announcement; E = event; R = report

Yellow highlight = infrastructures and initiatives in place

	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Leadership Council - National Summit																										
Community Input		X	X																							
DI Summit 2014			E																							
Policy Framework Agreed			X																							
Collective Action Plan/Roadmap Agreed			X	X																						
Report from Summit				X	X																					
Decision on Ongoing Coordination						D																				
TC3+																										
Consultations on Dig Scholarship Paper	X	X																								
Synthesis of consultations			X																							
Next Phase Policy Consultation						X	X																			
Policy Release on Data Mgmt									P	P																
CFI Workshop with Selected Researchers on DI Needs			E																							
CFI IF Competition				X	X	X	X	X	X	X	X	X	X	X	X	X	D									
CFI Cyber-infrastructure Competition					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	D						
NSERC/CIHR/GC/CFI Competition - Exploring Big Data	X	X	X	D																						
CIHR commissioned assessment by CCA of access to health data			X	X	X	X	X	X	X	X	X	X	X	X	???											
Compute Canada																										
Strategic Planning Consultations		X	X																							
Board Approval of Strategic Plan				D																						
Operational Planning & consults			X	X	X	X																				
Business Plan Approved					X	X	D																			
Optimization Plan					X	X	X																			
CFI Mid-term Review			TBD																							
Submission to CFI Cyber Competition			TBD																							
CANARIE																										
Strategic Plan Consultations & env scan	X	X	X	X																						
Issue Working Groups					X	X																				
Draft Strategic Pan; determine budget ask						X	X																			
Board review - draft Strategic Plan							X																			
Present draft strategic plan at CANHEIT								X																		
Board approval strategeic plan									D																	
Mandate renewal activities										X	X	X	X	X	X											

Research Data Canada																										
Research infrastructure Cttee		X	R	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Education & training Cttee		X	R	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Policy Cttee		X	X	R	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Standards & Interoperability Cttee			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Interface with international activities incl RDA			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
CARL																										
Launch of subgroup - Natl Data Network		X																								
Confirm charter and appoint project manager			X	X	X																					
Develop service model and business plan					X	X	X	X	X	X	X	X	X	X	Possibly ongoing											
Explore funding opportunities						X	X	X	X	X	X	X	X	X	Possibly ongoing											
CRKN																										
Develop - Integrated Digital Scholarship Ecosystem (IDSE) Project			X	X	X	X	X	X																		
Approval of IDSE scope											D															
Implementation of IDSE												X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Canadian Representative to/in SCOAP3		D	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
CASRAI																										
Research classification standards	X	X	X	X	X	X	X	X	D1	X	X	X	X	X	X	X	X	X	X	D2						
Res Admin standards for Data Mgmt Plans UK	X	X	X	X	X	X	X	X	X																	
Canada - Res Admin standards for Data Mgmt Plans																				X	X	X	X	X	X	X
Regional and Institutional Initiatives - examples																										
UBC-UofA-SFU consortium project on digital preservation	X	X	X	X	X	X																				
Archivematica-as-a-service (AAS) to COPPUL initiative			X	X	X	X	X	X	X	X	X	X	X	X	X	X										
Possible migrat'n of AAS as a preservation serviceX Canada													X	X	X	X	X	X	X	X	X	X	X	X	X	X
UBC-UofA-OCUL Cloud interoperability and data sharing																	X	X	X	X	X	X	X	X	X	X
OCUL Cloud storage initiative	X	X	X	X	X																					
OCUL Scholars Portal Initiatives - integrated res support	X	X	X	X	X	X	X	X	X	X	X	X	X	X												
Domain Specific DI Training amd DI Research Initiatives																										
Digital Humanities Workshop - Whistler				E																						
Interogating Access Conference - Wilfred Laurier				E																						
DHSI - Digital Humanities Summer Institute								X												X						
Statistical Learning for Big Data - Fields Institute Program															X	X	X	X	X	X						
Domain specific Digital Infrastructures																										
INKE - Implementing new Knowledge Environments	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
CWRC - The Canadian Writing Research Collaboratory	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
The Alberta Land Settlement Infrastructure Project																										
Canadian Century Research Infrastructure (CCRI)																										
Synergies																										
TAPoR and Voyant																										
Other Digital Infrastructure																										
Islandora																										
Public Knowledge Project																										

Brief Descriptions of Certain of the Targeted Digital Infrastructure Initiatives[3]

The CIHR Request for a CCA Assessment

CIHR has commissioned the Council of Canadian Academies to undertake and assessment of "timely access to health and social data for health research and health system innovation in Canada". The Council's assessment will examine the full range of barriers and challenges that limit access to health data, and assess evidence pertaining to risks and opportunities of "big data" research. The assessment will also consider best practices of Canadian and international health research institutions with respect to data management and governance.

Challenge question - What is the current state of knowledge surrounding timely access to health and social data for health research and health system innovation in Canada?

Creation of a National Research Data Network – Facilitated by CARL (and CRKN)

CARL is facilitating the creation of a collaborative network for collecting, preserving and providing access to valuable research data produced in Canada, a network that builds upon emerging initiatives and fills existing gaps. CRKN will join this initiative to provide the necessary breadth of engagement. A working group is being convened with financial support from CARL's program budget. The working group will include representatives of the research data management initiatives already in development in Alberta, British Columbia and Ontario, and representatives of other organizations interested in developing the national network, including Quebec and the Atlantic region. It is anticipated that this initiative will require a dedicated project manager with knowledge of the field. One of the first tasks of the group will be to identify and address such needs.

Key tasks for the group will be to develop a detailed service model and business plan that acknowledge the existing investment of the regional networks and provide opportunities for institutions outside those regions. These will outline the division of services and infrastructure across institutions, develop research data management collection policies and identify funding required to build capacity and sustain operations. Funding opportunities will be explored concurrently with other aspects of the planning, as a top priority. The group will also propose a model for ongoing management of the network. The working group will be dissolved once ongoing collaborative agreements and roles are in place.

Western Consortium

The University of Alberta Libraries, Simon Fraser University Library and the University of British Columbia Library are working together on a collaborative project to develop a research data management infrastructure that could serve as the regional preservation backbone in Western Canada, and could become a component of a national digital preservation network

[3] This list is indicative of some of the Canadian DI initiatives that exist or are under development, but is by no means comprehensive. There is a need to create a more comprehensive inventory of initiatives in health and the natural sciences and engineering and to foster interfaces among diverse fields of research.

Archivematica-as-a-Service
This initiative envisions providing Archivematica-as-a-service to COPPUL member organizations who wish to preserve digital holdings but who are unable or unwilling to install and manage local Archivematica instances. Archivematica is a free and open-source digital preservation system that is designed to maintain standards-based, long-term access to collections of digital objects (https://www.archivematica.org/).

Research Data Storage - OCUL
OCUL (Ontario Council of University Libraries) is developing a community cloud storage network to address members' data storage needs in a collaborative way. This is creating a critical element of the digital infrastructure necessary for national digital preservation and research data initiatives.

Statistical Learning for Big Data. Fields Institute of Mathematics/CANSSI Program proposed for *January – June 2015*
The topic in it broadest version is "Big Data" which refers to data that cannot be analyzed with conventional statistical and computational techniques because the size (terabytes and beyond) is too large to be manipulated with standard software and/or the models are too complex to allow computation of key quantities for inference. The more focused topics include the study and advancement of inferential techniques for statistical learning in big data. Two important goals of the program are to foster new collaborations across computer science, statistics and pure and applied mathematics, and to emphasize the relevance of statistical inference and the role of uncertainty quantification in analysis.

DHSI - The Digital Humanities Summer Institute at the University of Victoria
The DHSI provides an environment for discussing and learning about new computing technologies and how they are influencing teaching, research, dissemination, creation, and preservation in different disciplines, via a community-based approach.
During a week of intensive coursework, seminars, and lectures, participants share ideas and methods, and develop expertise in using advanced technologies. Every summer, the institute brings together faculty, staff, and students from the Arts, Humanities, Library, and Archives communities as well as independent scholars and participants from areas beyond.

The CWRC - The Canadian Writing Research Collaboratory / Le Collaboratoire scientifique des écrits du Canada

The CWRC is an online infrastructure project designed to enable unprecedented avenues for studying the words that most move people in and about Canada. It is designed to offer:

- a database (Online Research Canada, ORCA) to house born-digital scholarly materials, digitized texts, and metadata (indices, annotations, cross-references). Content and tools will be open access wherever possible and designed for interoperability with each other and with other systems..

- a toolkit for empowering new collaborative modes of scholarly writing online; editing, annotating, and analyzing materials in and beyond ORCA; discovering and collaborating with researchers with intersecting interests; mining knowledge about relations, events and trends, through automated methods and interactive visualizations; and analyzing the system's usage patterns to discover areas for further investigation.

INKE – Implementing New Knowledge Environments *(http://inke.ca/)*
INKE is a network that brings together researchers and stakeholders at the forefront of computing in the humanities, text analysis, information studies, usability and interface design into a network comprised of those who are best-poised to understand the nature of the human record as it intersects with the computer, with its work divided at present into three key research groupings: textual studies, modelling and prototyping and interface design.

Building Partnerships to Transform Scholarly Publishing (Digital Humanities)
This gathering, facilitated by Implementing New Knowledge Environments intends to provoke collaborations in the realm of scholarly communication and publication. The one and a half day gathering will provoke collaboration and conversation around electronic scholarly journals and monographs, as well as issues of (open) access, dissemination, alternative modes of scholarly communication, and the move from prototyping to producing.

Islandora
http://islandora.ca/
Islandora is an open-source software framework designed to help institutions and organizations and their audiences collaboratively manage, and discover digital assets using a best-practices framework. Islandora was originally developed by the University of Prince Edward Island's Robertson Library, but is now implemented and contributed to by an ever-growing international community.

The Last Best West: The Alberta Land Settlement Infrastructure Project(ALSIP)
http://www.abbeytoday.com/infrastructure/index.php
The Alberta Land Settlement Infrastructure Project (ALSIP) has three linked parts: digitized images of the Alberta Homestead Records; an on-line database compiled from those images, extended and enhanced by the addition of geo-physical variables; and 100 percent of Alberta's population in the 1911 census linked to the Homestead records.

Canada Century Research Infrastructure (CCRI)
http://ccri.library.ualberta.ca/enindex.html
The CCRI represents an infrastructure that facilitates research on the transformation of Canadian society in the twentieth century. The CCRI's mandate is to provide researchers with a body of data and information that can be used to acquire a better understanding of how modern-day Canada has developed. The CCRI is composed of microdata, namely, data created from Canadian census enumerations between 1911 to 1951, a geographical framework constructed to enable the location, selection, aggregation, and analysis of census data, and

contextual data, namely the textual data used to situate the census in time and to enhance appropriate analysis of the data.

TAPoR and Voyant

http://tapor.ca and http://voyant-tools.org

TAPoR and Voyant are two coordinated projects provide a) an environment for the discovery and review of text analysis tools, and b) an online text analysis environment. They are used over 40,000 times a month and are probably the most used tools of their kind in the humanities.

Public Knowledge Project

PKP is a multi-university initiative developing (free) open source software and conducting research to improve the quality and reach of scholarly publishing. It has developed a number of important open systems:

- Open Journal Systems (OJS) is a journal management and publishing system that has been developed by the Public Knowledge Project through its federally funded efforts to expand and improve access to research.
- Open Monograph Press is an open source software platform for managing the editorial workflow required to see monographs, edited volumes and, scholarly editions through internal and external review, editing, cataloguing, production, and publication. OMP can operate, as well, as a press website with catalog, distribution, and sales capacities.
- Open Conference Systems (OCS) is a free Web publishing tool that will create a complete Web presence for your scholarly conference.
- The Open Harvester Systems is a free metadata indexing system developed by the Public Knowledge Project through its federally funded efforts to expand and improve access to research.

Synergies

http://www.synergiescanada.org/

Synergies : Canada's SSH Research Infrastructure - A not-for-profit platform for the publication and the dissemination of research results in social sciences and humanities published in Canada.

Appendix 2 – Stakeholders in the Canadian DI Ecosystem

Extracted from the TC3+ Consultation document (without modest updating) are brief descriptions of the various organizations and working groups active in promoting and supporting digital scholarship and the cyber-infrastructure underpinning in Canada.

Research Data Canada (RDC)

- **What** – A stakeholder-driven and supported national body dedicated to advancing the vision for research data in Canada.
- **Who** – A variety of stakeholder organizations, all with an interest and role to play in ensuring that the infrastructure, processes and support are in place to realize the vision for research data in Canada. This includes CUCCIO, Compute Canada, CANARIE, CARL, CFI, CIHR, NSERC, SSHRC, CASRAI, NRC, TB, IPY, LAC and CODATA.
- **Mandate** – To develop strategy, facilitate communication and partnerships among data initiatives, promote education and training in data skills, measure progress in implementing the vision, bring attention to gaps, and act as a single point of contact for Canada in international data initiatives (self-generated mandate).
- **Functions** – Its activities focus on five areas: policies, infrastructure, standards and interoperability, education and training, and international liaison. RDC does not plan to own or operate infrastructure.
- **Intersects/interfaces** – May be somewhat different from the Digital Infrastructure Leadership Council in that it focuses on issues related to research data and data lifecycle management. However, there does appear to be some overlap between the two in focus and participation.
- **Recent actions** – The inaugural RDC meeting was held in January 2013. RDC's priorities for 2013 are to:
  - **launch Research Data Canada** as the multi-stakeholder/volunteer-driven organization that will drive efforts forward to ensure that the full value of Canada's research data is realized;
  - **encourage broad membership** for Research Data Canada to reflect fully the diversity of stakeholders with an interest in research data;
  - **advance the work of the RDC Committees**: Infrastructure, Education and Training, Policy, Standards and Interoperability, and International Liaison;
  - **host Webinar series** on data management – Canadian and international speakers, range of topics, audiences;
  - **co-sponsor data stream in CASRAI Big Data Reconnect conference** October 2013 and coordinate a pre-conference Data Centres workshop;
  - **continue international liaison** with Research Data Alliance, Global Research Data Infrastructure, DataCite Federation and other initiatives;
  - **establish a national advisory council** of senior representatives from industry, the academy, government research labs, funding agencies and policymakers to provide counsel to Research Data Canada; and
  - **initiate a national online consultation** process to take the results of the 2011 Canadian Research Data Summit to a broader set of stakeholders across the country.
- **History:**

  - Took over from the Research Data Strategy Working Group (RDSWG) in 2012; originally formed to survey and identify the challenges and issues surrounding access to, and preservation of, data arising from Canadian research;
  - The RDSWG organized the September *2011 Canadian Research Data Summit: Mapping the Data Landscape*;
  - The final report of the Summit proposed a National Strategy for Research Data in Canada, including a vision statement, high-level goals and a framework for action with broad timelines and distribution of tasks across major stakeholder communities; and
  - Also conducted a gap analysis for Canada: http://publications.gc.ca/collections/collection_2009/cnrc-nrc/NR16-123-2008E.pdf
- The RDSWG transferred its activities to Research Data Canada in December 2012.

Leadership Council for Digital Infrastructure

- **What** – A stakeholder-driven and supported national initiative dedicated to developing a national strategy to renew and strengthen our current advanced digital infrastructure for research, innovation and education in Canada.
- **Who** – A cross-sector group co-chaired by Jay Black, SFU and CUCCIO, and Steven Liss, Vice Principal Research, Queen's University, with membership including: CRKN, CUCCIO, Compute Canada, CANARIE, CARL, CFI, NRC, Industry Canada, CIHR, NSERC, SSHRC and CASRAI.
- **Mandate** – To build on the work accomplished at the Summit, develop plans for first initiatives, and put in place the mechanism to ensure continued engagement of the many stakeholders.
- **Functions and priorities** – Conduct a gap analysis, develop a roadmap to address the gaps and convene a follow-up to the 2012 Summit.
- **Intersects/interfaces** – Different from Research Data Canada in that its focus is to provide an overarching view and the required mechanism(s) to support an integrated and sustainable approach to Canada's advanced digital infrastructure eco-system.
- **Recent actions** – Were a result of the CUCCIO-hosted Digital Infrastructure Summit 2012 in Saskatoon.

Canadian Association of Research Libraries (CARL)

- **What** – The leadership organization for the Canadian research library community.
- **Who** – Members are Canada's 31 large research libraries.
- **Mandate** – Provides leadership on behalf of Canada's research libraries and enhances their capacity to advance research and higher education. It promotes effective and sustainable scholarly communication, and public policy that enables broad access to scholarly information.
- **Functions and priorities** – Its *2013-2016 Strategic Directions* include the following points:
  - facilitate collaborations to share and preserve Canada's research collections;
  - coordinate research data management initiatives; and
  - promote open access and new forms of scholarly communication.
- **Recent actions related to digital scholarship**:
  - CARL led the development of a concept of a distributed network of data repositories, which was to inform a 2012 proposal for CFI funding; ultimately, a formal proposal was not submitted to CFI as the LEF-NIF was not an ideal CFI program for it.

- Along with CRKN, CARL published in October 2012 a report emphasizing the potential role(s) for academic libraries in implementing an open access policy on research publications.
- In January 2013, CARL held a four-day "Introduction to Research Data Management" course for librarians; about 60 individuals enrolled from 30 universities; these participants have continued to confer as an online "community of practice."
- CARL has worked to facilitate the development of open access digital repositories at (by now) all member libraries; most CARL member libraries are developing local research data management services for researchers.
- CARL has advocated for federal government support for national research data management infrastructure; it has developed web content, conference presentations, workshops and articles on both open access and research data management.
- CARL is working with CASRAI and Research Data Canada to produce CASRAI's 2013 Big Data conference.
- CARL has created a Data Management Subcommittee headed by its Vice-President/President-Elect (Martha Whitehead, Queen's University) that is proposing the formation of a national collaborative network of local/regional/other initiatives for collecting, preserving and providing access to research data produced in Canada.

- **Intersects/interfaces** – Both CRKN and Canadiana.org (Canada's premier organization for the digitization of Canadian historical documentation and the exposure of Canadian digital documentary collections) had their origins as CARL initiatives; CARL collaborates closely with them. CARL is a supporting member of Research Data Canada and the Leadership Council for Digital Infrastructure. CARL is the major Canadian association member of SPARC (the Scholarly Publishing and Academic Resources Coalition) and COAR (the Coalition of Open Access Repositories), both of which promote and support open access and repositories internationally. Locally, CARL member library directors work in consultation with CUCCIO member CIOs in the context of various digital services.

Ontario Council of University Libraries (OCUL)

- **What** – A library consortium
- **Who** – Ontario's 21 university libraries
- **Mandate** – To enhance information services in Ontario and beyond through collective purchasing and shared digital information infrastructure, collaborative planning, advocacy, assessment, research, partnerships, communications and professional development.
- **Functions** – Provide access to a diversity of learning and research materials, and ensure their preservation through sustainable and responsible stewardship; lead in the development of partnerships to expand Canada's digital research infrastructure; operate Scholars Portal; encourage the advancement of access to electronic data resources including those provided under the Data Liberation Initiative (DLI); expand access to maps, geospatial data and other cartographically related resources, both print and digital.
- **Recent activities –** Scholars Portal has received certification as the first Trustworthy Digital Repository in Canada. This certification, the only generally recognized certification for digital archives, was issued by the Center for Research Libraries (CRL).

Canadian University Council of Chief Information Officers (CUCCIO)

- **What** – CUCCIO is a non-profit, member-funded association of Canada's higher education information technology leaders, working together to help Canadian universities excel through the innovative and effective use of IT.
- **Who** – Composed of the chief information officers (CIO) from more than 50 universities Canada-wide.
- **Strategic priorities/mandate** (from Strategic Plan):
  - foster best practices in information technology management in Canadian universities;
  - identify, incubate and sponsor collaborative sector-wide services;
  - develop and deliver programs and services to support the professional development of IT staff; and
  - develop and maintain relationships with governments, government agencies, corporations and other groups of interest to higher education in order to advance the shared interests of Canadian universities.
- **Recent actions related to digital scholarship:**
  - Convened the *Digital Infrastructure Summit 2012* to:
    - establish a vision for a comprehensive, integrated and sustainable digital infrastructure in support of research, education and innovation in Canada;
    - develop a specific action plan, with milestones; and
    - secure commitments from stakeholders with regards to realizing the plan.
  - Emanating from the Summit, it created a cross-sector group, the *Leadership Council for Digital Infrastructure*, to build on the work accomplished at the Summit (above).
  - Plans are underway for a follow-up to Summit 2012.
- **Intersects/interfaces** – The work of the Council will intersect with the strategic and operational plans of Compute Canada, CANARIE and Research Data Canada from the perspective of the digital infrastructure overall and to facilitate as possible integration and coordination of the various components.

Compute Canada

- **What** – An incorporated NFP organization that provides Canada's national platform of High Performance Computing (HPC) resources. A new President and Board are in place effective late 2012.
- **Who –** Members are 29 Canadian universities. Membership is available to any university or college in Canada that has one or more researchers using an advanced computing system, through access provided by Compute Canada. Compute Canada assigns services on a pan-Canadian basis, using regional nodes:
  - Compute West – WestGrid
  - Compute Ontario – HPCL, SciNet, SHARCNet
  - Calcul Quebec – formerly RQCHP and CLUMEQ
  - Compute Atlantic – ACENet
- **Functions –** Compute Canada delivers its services through HPC systems managed by regional consortia at different locations across Canada and utilizing the CANARIE broadband network.
- **Mandate** – To promote and support the shared use of advanced computing resources designed to keep Canada competitive in research and innovation. It is not, however, a policy-making body or a standards organization.

- **Assets** – The Compute Canada platform includes computing capability, online and long-term storage, connection to the CANARIE network, and user support services. It is primarily oriented toward larger computation systems used in simulation and computational intensive research.
- **Funding** – Capital assets funded in part by CFI. Institutions also a player. Also in receipt of significant support for operating and maintenance from the CFI MSI Fund.
- **Recent and current actions**:
- Recent unsuccessful proposal to CFI (LEF-NIF competition) to address the needs of the medical and SS&H communities for:
  - Issue at the time was more capacity than direction (which was deemed good).
  - Development of a strategic plan for the organization through 2013.
- **Intersects/interfaces –** Provinces, CANARIE, RDC, Leadership Council, private sector.

CANARIE

- **What** – CANARIE supports research and education through the delivery of advanced digital infrastructure. Manages and evolves one of the world's largest and fastest research and education networks, in partnership with provincial and territorial networks (ORANs).
- **Who** – A not-for-profit corporation funded primarily through the federal government, with additional funding from membership revenues and fees for services.
- **Mandate** – to design and deliver digital infrastructure, and drive its adoption for Canada's research, education and innovation.
- **Functions** – manages an ultra-high-speed national backbone network that connects provincial and territorial networks to each other and to 100 international networks. The provincial and territorial networks connect directly to universities, research centers, government labs, hospitals and other scientific facilities within their jurisdictions, and to CANARIE's national backbone and global research and education networks. The partnership of CANARIE and the ORANs enables researchers, educators and innovators to move, share and analyze data and access specialized tools and resources. CANARIE also supports the development of software to support research collaboration and access to widely distributed data and tools. CANARIE spurs research and innovation in the private sector by offering small and medium-sized businesses access to a cloud-based testbed to accelerate product development timelines and reduce costs.
- **Objectives** for the period April 1, 2012 to March 31, 2015 are:
  - **Network Operations –** Operation and evolution of the CANARIE network as essential research infrastructure; extending the "owned" portions of the network to provide greater flexibility and lower costs for delivering greater bandwidth, including network-based services, which currently include:
    - Canadian Access Federation (CAF);
    - Content Delivery Service (CDS); and
    - managing the Network Alliance Infrastructure and Network Alliance Development programs to strengthen the pan-Canadian network and enhance the visibility of this essential digital infrastructure.
  - **Technology Innovation** – Develop, demonstrate and implement next-generation technologies to advance the CANARIE network as a leading-edge research network—including new software tools, comprising a toolkit of reusable services. Two programs support this objective:

        - Research Platform Interfaces (RPI) **–** Leverages services from the previous Network-Enabled Platforms (NEP) program by creating a collection of platform services (RPIs) from existing NEPs to be used by multiple research platforms; and
        - Network-Enabled Platforms (NEP) **–** The development of sophisticated software platforms that enable researchers to easily collaborate and access research data and tools.
    - **Private Sector Innovation** – Leveraging the CANARIE network to assist firms operating in Canada, and Canadian universities, to advance innovation and commercialization of products and services to bolster Canada's technology innovation capabilities. CANARIE's DAIR (Digital Accelerator for Innovation and Research) program offers a cloud-computing testbed for small and medium-sized enterprises to accelerate product development, reduce costs and realize the scale and agility benefits of cloud technologies.

ORANs

- **What** – Optical Regional Advanced Networks (ORANs)
- **Who** – 12 regional networks:
    - East – ACORN-NL, ACORN-NS, New Brunswick Advanced Network, Prince Edward Island Advanced Network;
    - Quebec – RISQ;
    - Ontario – ORION;
    - West – BCNet, Cybera, SRnet, MRnet; and
    - Territories – Aurora College, Yukon College.
- **Mandate** – To support the operation and development of advanced networks and services at the regional level in support of research and innovation.
- **Intersects/interfaces** – The ORANs provide connectivity of the regional high-speed network to the national backbone provided by CANARIE.

Canadian Access Federation

- **What** – A trusted access management environment (single-identity access) for Canadian research and higher education communities.
- **Who** – Developed by CUCCIO with operational responsibility transferred to CANARIE in 2012.
- **Mandate** – To make sharing protected resources easier, safer and more scalable by:
    - enabling staff, students and faculty to access wireless networks and web-based resources using their home organization credentials when they are visiting other organizations;
    - allowing participants to participate in a cost-effective, privacy-preserving approach to access management;
    - helping to ensure the privacy of personal information by eliminating the need for researchers, students and educators to maintain multiple, password-protected accounts; and
    - enabling organizations to better manage access to their resources based on a user's status and privileges as presented by the user's home organization.

Canadian Research Knowledge Network (CRKN)

- ***What*** – An organization of Canadian universities dedicated to expanding digital content for the academic research enterprise in Canada.

- **Who** – An incorporated NFP organization that is a partnership of 75 Canadian universities. University libraries are the drivers of CRKN's initiatives, and play a primary role in leveraging expertise and resources for the benefit of Canada's scholarly research community.
- **Mandate** – To undertake large-scale content acquisition and licensing initiatives in order to build knowledge infrastructure and research capacity in Canada's universities and to provide equitable and cost-effective access to scholarly content for universities nationwide.
- **Recent activities** – CRKN's current draft strategic plan identifies a role for the organization in supporting and coordinating vertical integration of all types of research data and digital scholarship nationally.
- **Intersects/interfaces** – Involves all of the AUCC libraries including the members of CARL plus approximately 40 others. CRKN has worked with CUCCIO to develop the Canadian Access Federation, and collaborates with regional academic library consortia including COPPUL, OCUL, CREPUQ and CAUL. CRKN represents Canada on the SCOAP3 Open Access initiative.

NRC Knowledge Management (formerly Canada Institute for Scientific and Technical Information or CISTI)

- **What:** Canada's national science library and NRC knowledge management organization
- **Who -** Division of NRC
- **Mandate –** Has made a major transformation since 2010 from science journal publisher and worldwide delivery hub for STM information, to knowledge and information services for NRC and the innovation community that include NRC Foresight and competitive and technical intelligence, to information management.
- **Recent Activities:**
  - Built a mirror site for NIH PubMed Central, in collaboration with COHR and the US Library of Medicine. PMC Canada provides Canadian researchers with access to American publications in PubMed Central and provides a portal for the deposit of Canadian journal articles. The portal supports CIHR's open access policy.
  - Developed and manages DataCite Canada, to manage data registration services, as member of the International DataCite Federation
  - Provides Canadian representation to the G8+06 Data Management WG
- **Intersects/interfaces** – CISTI has also acted as host and administrator for various data and open access working groups, including the Research Data Canada since 2008 —activities that have linked it with the funding agencies, CARL, CASRAI, the CRDCN, etc.

CASRAI

- **What** – A not-for-profit standards development organization focusing on research administration data. The Board represents the diversity of stakeholders.
- **Who** – A community of research organizations (funders—federal and provincial, institutions, implementers) collaborating to evolve the standard dictionary of research terminology and to advance the standard platform for research interoperability. Representatives from participating organizations sit on a number of committees, review circles and advisory councils. There are international mirrors of CASRAI Canada developing though the leadership of CASRAI.

- **Mandate** – To provide a forum and the mechanisms required to standardize the data that researchers, their institutions and their funders must produce, store, exchange and process throughout the life-cycle of research activity.
- **Priorities** – To advance those semantic standards that will facilitate effective operation in the digital business environment.
- **Intersects/interfaces** – Participates in the Leadership Council and the RDC.

Canadian Research Data Centre Network (CRDCN)

- **What –** The Network acts as a pan-Canadian forum and structure to give Canada's research community access to social and population health statistics and help provide evidence for effective public policy and planning. CRDCN oversees a pan-Canadian array of Research Data Centres (RDCs). An RDC is a university-based laboratory, staffed by a Statistics Canada Analyst, that offers researchers on-site services for:
  - **Secure access to confidential micro-data** – Statistics Canada census and surveys, plus a growing range of administrative data; prospects of datasets from other federal departments.
  - **What they need to analyze the data** – Fully-equipped workstations, statistical software and technical support
- **Who –** 45 academic institutions and Statistics Canada form the core membership of the Network.
- **Mandate:**
  - **To improve data access** by giving researchers across the country access, free-of-charge, to detailed micro-data from an increasing range of survey, census and administrative data.
  - **To expand the pool of skilled quantitative researchers** in Canada and train the next generation of researchers.
  - **To make research count** by improving communication between social scientists and the potential users of the knowledge they create.
- **Intersects/interfaces** – Statistics Canada, CANARIE, the ORANs, NRC-CISTI, various social and health policy federal departments and agencies.

Canadian Polar Data Network

- **What** – A Canadian network and standards-based organization that grew out of the data centre for the International Polar Year.
- **Who** – A partnership of the University of Alberta Libraries, University of Waterloo Canadian Cryospheric Information Network, OCUL Scholars Portal, Fisheries and Oceans Canada – Integrated Science Data Management, and NRC-CISTI.
- **Mandate** – To provide a sustainable research data management infrastructure, encompassing preservation and access, for polar (Arctic and Antarctic) science research and monitoring initiated from and taking place in Canada.
- **Intersects/interfaces** – Those government departments and agencies that have been designated by federal legislative documents (for example, the *Oceans Act*) to collect data for the purpose of understanding the environment and its living resources and ecosystems.
- **Priorities** – Data in scope but outside the Government data archival divisions will be a priority for the CPDN.

Canadian Astronomy Data Centre (CADC)

- **What** – One of the principal data archiving and data mining facilities worldwide for astronomical data.
- **Who** – A division of NRC.
- **Mandate** – Management, curation, preservation and access of all data for projects in which Canadian astronomers are involved (primarily academic).
- **Intersects** – The international astronomical community; the Canadian Space Agency.