



A Policy Framework for Canadian Digital Infrastructure¹

Introduction and Context

The Canadian advanced digital infrastructure (DI) ecosystem is the facilities, services and capacities that provide the research community (academic, government and private sector) with the capability to conduct data-intensive, top-level research in their respective fields, and to intensify collaborations nationally and internationally. The research infrastructure includes single-site, distributed and nationally centralized e-infrastructure that is itself connected with comparable systems internationally. That DI ecosystem increasingly provides access to data resources for the private sector and Canadians at large. A further dimension that defines the particular public-private partnerships required to support, and generate, larger societal value from, the DI ecosystem will be a cornerstone of Canada's approach to digital infrastructure.

This policy framework (and associated roadmap) provides the context in which the DI ecosystem is funded and managed for the benefit of all Canadians. It is a living document that drives Canada's global leadership and will continue to evolve on a regular basis through periodic review and assessment against key metrics or benchmarks.

While particularly critical for the research sector, the benefits of investment in an effective DI ecosystem will soon permeate all sectors of the economy and society, as described below.

Benefits for Canadians of a Robust and Sustainable DI Ecosystem

- **Enabling Canadian researchers to compete and collaborate globally** – Big Data and computational and transmission power are changing the way research is done (e.g. “in silico” science), changing the pace of discovery and opening up entirely new fields.
- **Supporting the training of highly skilled personnel** – the DI ecosystem enables the training necessary to create a new generation of data scientists – people who can assume pivotal roles in data-intensive careers in private and public sectors.
- **Enabling the sharing and re-use of data** – clear and efficient arrangements for the exchange of data allows researchers and other users from the private and public sectors and the lay public to access primary data for diverse purposes.

¹ Drafted for DI Summit 2014 by the Project consulting team.

- **Fostering innovation and catalysis of private sector growth opportunities** – access to highly skilled personnel and public data is a source of increased efficiency and effectiveness, new product lines, new public-private partnerships, new firms and industrial clusters, and new markets.
- **Providing an enriched base for government decision making** – governments can draw on and assess an enriched database of knowledge and understanding for policy analysis and decision making as well as an enhanced capacity to monitor the impact of prior decisions, as well as making services and resources more available to the public.
- **Enabling citizen science** – the lay public can access and utilize, and sometimes contribute to, the huge amount of public data available for personal knowledge and entertainment and for contributing to research projects.
- **Supporting productive connections** – using data as the new currency, the DI ecosystem will catalyze increased linkages among Canadian researchers and data users and within and across sectors, and across nations.

The following text boxes articulate the vision, principles and objectives of the Canadian DI ecosystem, as initially captured at the 2012 DI Summit, and also the mission and goals of the Leadership Council for Digital Infrastructure.

The Digital Infrastructure Ecosystem in Canada

Vision:

Canada’s world-class digital infrastructure ecosystem benefits society by accelerating research, innovation and education.

Principles:

- Integrated:** All essential elements are integrated in a co-ordinated national system.
- Inclusive:** The national system accommodates users from a full range of disciplines and sectors.
- Comprehensive:** The national system provides a full spectrum of service.
- Accessible:** The national system is accessible to users regardless of their location, discipline, sector, or level of expertise.
- Valued:** The national system receives attention and is given priority.
- Governed:** The national system has effective oversight, control and direction.
- Sustainable:** The national system is designed and funded so that it can evolve and adapt.

The Leadership Council for Digital Infrastructure

Mission:

To foster a DI ecosystem in which researchers and other end-users have rational, effective, agile and

efficient access to tools and resources to enable world class research, research training and innovation within and across a wide range of disciplines and sectors.

Goals

To provide:

- A platform for all national DI stakeholders to collaborate in developing a sustainable framework for digital infrastructure for research, education and innovation in Canada.
- A forum to identify, discuss and address the issues associated with providing Canadian researchers and other end-users with rational, agile and efficient access to the tools and resources they require to enable research within and across a wide range of disciplines and sectors.
- A catalyst for the development of the initial conditions required to implement and operationalize the DI in a practical and timely manner in Canada.

The Policy Framework

The Purpose of the Framework

The purpose of this Policy Framework is to build on the above vision and principles and to address:

- **Coordination:** foster active and effective liaison, alignment, and coordination among the institutions and organizations that share responsibility for delivering a robust and sustainable ecosystem for digital infrastructure;
- **Roles and responsibilities:** provide clarity on the individual and shared roles and responsibilities of the various players in the development, provision, funding and management of digital infrastructure;
- **Benefits:** describe the conditions under which digital infrastructure will deliver optimal outcomes for research and larger socio-economic benefits;
- **Leadership:** provide leadership in implementing and operationalizing the DI; and
- **Funding:** articulate the principles that guide funding decisions on digital infrastructure for research.

The Players

See Appendix 1 for a list of the key players in the Canadian DI ecosystem. The list is long and diverse; hence the very significant need for an effective means of collaboration and coherence in planning and implementation.

Principles Underlying the Framework

Liaison, Coordination and Alignment

1. **Coordinated strategic planning** – A quality DI ecosystem requires planning at each level of the research system: i) within disciplines with a high level of data and computational intensity (e.g. particle physics); ii) across research communities that can benefit from common data analytics and shared approaches to data management; iii) within and among professional associations (e.g. CARL),

and iv) among funding agencies. That planning is only valuable, however, if communicated and acted upon.

2. **Coordinated service delivery planning** – CANARIE, the ORANS, Compute Canada and Research Data Canada and CARL (representing the research data management players) will ensure alignment and complementarity of their strategic and operational plans.
3. **Coordinated investment** - A robust and sustainable pan-Canadian DI requires ongoing, coordinated investment from funders and institutions.
4. **Coordinated investment, oversight and balance** – Working together, the federal funders (CFI and Industry Canada) will review and maintain a suitable balance in funding among the three pillars of the national DI ecosystem – the national broadband network, advanced computational capacity and research data management infrastructure, as well as a suitable balance between capital and operating costs.

Leadership

5. **Shared leadership** entails shared responsibility and accountability, openness and transparency, mutual trust, a commitment to consensus building, and the use of evidence in decision making.
6. **Keeping the policy framework and roadmap evergreen** - This policy framework and roadmap will continue to evolve under the auspices of the Leadership Council to ensure that they are kept current and relevant to the evolving needs of all stakeholders in the DI ecosystem.

Roles and Responsibilities

7. **Framework conditions** - Policy and funding agencies have the responsibility to set the framework conditions for an effective DI ecosystem, in particular for the management of data resources, the sharing of advanced computational facilities, the conditions for maintenance and utilization of national broadband networks, the expectations for coordination among the key players, and high level review regarding overall DI integration, functionality, performance, and achievements.
8. **Co-investment** - Co-investment in DI, especially at regional and local levels, is required for Canadian researchers to have access to the necessary level of resources. It also demonstrates a commitment by the investing party/ies to the health of the DI ecosystem. Co-investment requirements will be flexible, however, and reflect the degree of local/regional returns on that investment.
9. **National investment for national facilities** - Within the context of shared funding for the overall DI ecosystem – both capital and operating – there will be a significantly higher level of national investment for pan-Canadian initiatives; more co-investment is expected for regional and local initiatives.
10. **Institutions and their libraries** - Academic libraries increasingly serve as the campus hub for accessing information, including research data, for researchers and students. They are also service providers to faculty and graduate students in the use and reuse of data, training, helping them

comply with requirements, and providing analysis facilities. This is a transformation of the notion of an academic library. As such, institutional libraries will assume responsibility for institutional data repositories provided there is an institutional commitment to providing the resources for repurposing of facilities and training/retraining of staff. Those repositories do not, however, fulfill the need for national repositories that must be developed in tandem with institutional repositories.

11. **National data repositories** – federal funding agencies and those government departments conducting research bear a particular responsibility to facilitate the evolution and inter-operability of national and regional data repositories (including those derived from large-scale discipline-specific investments) that complement the institutional repository system.
12. **Private sector engagement** – Compute Canada, CANARIE and other agencies responsible for advanced digital infrastructure will foster entrepreneurialism and explore expanded relationships with the private sector – as entities at the cutting edge of advanced hardware and software, as users of the facilities and as collaborators in frontier research initiatives.

Optimizing Benefits

13. A **culture of stewardship** of data resources and the associated computational capacity for community benefit will replace that of individual ownership. This will entail mutual trust and shared benefit, supported by suitable incentives for compliance and consequences for non-compliance.
14. **Funding duration** will be aligned with the time-scales appropriate for large scale infrastructure that serves national needs, with realistic and predictable timelines for review and renewal of support.
15. **Data will be included in the definition of digital infrastructure** when available for use and reuse. Data are a national resource that require professional management and curation, effective handling of access provisions, protection of privacy and confidentiality, and long-term preservation when justified by the value of the resource.
16. **Data as a resource for research** - As data are of increasing importance as a resource for research (not just a product), there will be value in preserving at least some data well beyond the career of a researcher or what might reasonably be expected of an institution's retention.
17. **Skilled personnel will be included in the definition of digital infrastructure** when providing support to research users of the national DI. It will be recognized that the real challenge in optimizing return on investment in DI is human, not technological.
18. **The benefits and costs of collaboration** among entities delivering services, policies and funding for DI will be acknowledged through appropriate funding mechanisms.

Guiding Investment Decisions

19. **Holistic** – DI investments will be made in the context of a broad interpretation of DI, including requirements for skilled personnel, training, user support, data storage, research data management systems, high speed networks, advanced computational capacity, tools and middleware, and

interoperability. Imbalance in investment among these components will compromise the functionality of the overall system.

20. **Sustainable** – National DI will be funded and managed as a “persistent infrastructure” with the stability essential for enabling frontier research, innovation and international connections, as well as ensuring long-term curation and preservation of valuable data. Stable and sustained funding is a necessity for the key infrastructure entities to acquire and maintain the capacities and expertise required to gain optimal benefit from the capital investments. A rigorous review process, with “institutional memory” that monitors progress over time, will provide assurance that DI entities meet performance expectations.
21. **Incentive structures** - Funding mechanisms will provide substantial positive incentives for inter-institutional, regional and national investment in sustainable and globally connected DI, and for individual researchers and end-users to comply with data management policies.
22. **Efficient** - DI will be delivered and funded at national, regional and local levels according to the most effective and efficient modes of delivery of the services, including the provision of user support.
23. **Flexible and future-oriented** – DI investments will foster flexibility, diversity of users and evolution with changing research demands, technologies, best practices and user needs.
24. **Integrated and inter-operable** – Canadian DI will support the sharing of distributed and heterogeneous computational and data resources across disciplines, sectors, regions, and nations. Effective standards (semantic and technical) for research data and research administrative data underpin this capability.
25. **Quality** – DI investments are assessed in a flexible manner responding to their scale on the basis of i) their importance for supporting excellent research; ii) the scientific and technological excellence of the DI; iii) the robustness of the governance and operating structures relative to the scale of the DI; iv) the nature and degree of user support; v) importance within the national DI ecosystem; and vi) the scale and importance of Canadian benefits that will result. Larger investments will require the most attention to points iii), iv), v), and (especially) vi).
26. **Predictable** – Cost-effective pan-Canadian DI investments require a long-term planning horizon and predictable mechanisms for accessing ongoing support. A minimum 5-year rolling planning horizon is essential.
27. **Open** – Pan-Canadian DI investments require associated policies for open access wherever possible: open-source code, open data and open courseware—with terms of use that encourage reuse.

Version 3.6
November 24, 2013

Appendix 1 – The Players

The DI ecosystem incorporates a large number of players as below:

Users

- Researchers from academe, government and the private sector
- Institutions, including hospitals and research platforms that deal with “big data”
- Educational institutions using the broadband networks

Institutional representatives of users’ interests

- Vice Presidents Research
- Chief Information Officers
- University Librarians / Deans of Libraries

Service Providers

- Compute Canada
- CANARIE
- The provincial and territorial ORANS
- Private sector providers of services and technologies
- University libraries
- The National Research Council for DataCite
- Research platforms that provide sophisticated software, data storage and data format services for specific applications (e.g. TRIUMF and Ocean Networks Canada)
- The Canadian Research Data Centre Network (providing confidential access to large microdata sets from Statistics Canada and in some cases provincial statistical offices)
- Government departments and ministries that collect and/or manage datasets for administrative and research purposes (e.g. DFO and NRCan; provincial health authorities)
- The Canadian Research Knowledge Network (CRKN)
- CASRAI – providing standards for research administration data

Allied Organizations and Associations

- Research Data Canada
- Various national and international organizations that forecast research requirements and/or establish standards for research data (often disciplinary in nature)
- Canadian Association of Research Libraries (CARL)
- Canadian University Council of Chief Information Officers (CUCCIO)
- Association of Universities and Colleges of Canada (AUCC)
- Polytechnics Canada (PC)
- Association of Canadian Community Colleges (ACCC)

Funders

- The federal government through Industry Canada
- The Canada Foundation for Innovation (CFI)
- The Tri-Council (Natural Sciences and Engineering Research Council NSERC, the Social Sciences and Humanities Research Council SSHRC and the Canadian Institutes of Health Research CIHR)
- Genome Canada and its regional centres
- The provinces and territories
- The private sector