"Think Piece" on a DI Roadmap[1]

PREAMBLE

This approach to a Digital Infrastructure (DI) Roadmap has been prepared on the basis of extensive community discussion and consultation. It suggests some of the more important steps to take to move towards achieving a robust and sustainable digital infrastructure for research in Canada. These steps reflect both the strong attributes of the digital infrastructure ecosystem that exist today (and on which the proposed actions are built), as well as the particular problems identified by stakeholders and the forward challenges of adapting to the rapidly evolving needs of data-intensive research in a time of fiscal restraint. At the same time, the draft document is just that – a draft "think piece" designed to trigger focused discussion on how best to shape our initial and mid-term actions. It is an example of the type of planning and actions required, including designation of roles and responsibilities, setting targets and time lines, and providing a means of coordination and assessment of progress.

Following DI Summit 2014, the Leadership Council will prepare a version of the Roadmap based on the input received and then launch or trigger action on the key priority actions.

Participants in the DI Summit and those providing comment through the website are encouraged to use this document to facilitate i) their advance thinking on what they would like to achieve and ii) their at-the-Summit discussion of priority actions. Among the questions that participants may wish to reflect on are:

---

[1] This Roadmap has been prepared by the Project consulting team and does not pretend to be representative of the views of all organizations participating in the Leadership Council.

Advancing Canada's digital infrastructure //Améliorer l'infrastructure numérique du Canada

1. If you agree with the need for some form of multi-stakeholder coordination and/or a "Coalition/Leadership Council" type of entity, how would you see it structured and working? How should it relate to any bodies created by the TC3+?
2. Are the proposed priority thrusts and approaches the best ones to make a difference? Are sufficiently fundamental changes being examined?
3. What three changes over the next two years would most improve the quality and sustainability of DI in Canada?
4. Recognizing funding constraints, are the proposed refinements to the planning and funding system appropriate? What would you propose if different from those identified?
5. How can we ensure the engagement of university officials – Presidents, Vice-Presidents Research, CIOs, librarians, Research Offices - in developing a more comprehensive and effective approach to the management of research data?
6. What can you and/or your organization do to improve Canadian DI?

Contents

1. OVERVIEW

This draft Roadmap has been developed to chart the next steps necessary to move towards a robust and sustainable DI ecosystem. It is the product of extensive discussions with research users, funders, service providers and other stakeholders and will be reviewed and revised at Summit 2014 to ensure relevance, feasibility and buy-in. Comments via e-mail from members of the research community who will not be involved in Summit 2014 are also welcomed.

Advancing Canada's digital infrastructure //Améliorer l'infrastructure numérique du Canada

2. WHAT IS THE PROBLEM?

Proposing action items requires significant commitment of all stakeholders and a clear definition of what problem we are trying to solve. The following, in a very general way, captures the essence of the major challenges we are facing.

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.

This paper articulates actions proposed to address the above problems – suggestions that are not cast in stone, but are advanced as a jumping off point for further discussions and exploration. There is not, however, a neat and tidy way to map solutions to each of the problem areas individually because of the interdependencies among the problems and potential solutions to each. The reader of this document should therefore asses how the overall portfolio of actions addresses the defined problem areas and what alternative or additional actions are required.

3. THE ACTION AGENDA

Digital infrastructure is fundamental to research, education and innovation in Canada. While Canada has made significant investments in digital infrastructure, there is significant opportunity to optimize these investments to drive greater benefits to Canada. In particular, there are opportunities to optimize DI planning, funding structures, investment decisions, user access, and timeliness of actions. This would mainly be achieved through greater alignment among the mandates and operations of service providers, and through development of a long-term strategy for Canada's significant data assets.

This Roadmap lays out proposals for methods to align the various organizational structures active in the DI ecosystem, clarify their relative roles and responsibilities, and reassess funding requirements. It does not propose a new service delivery or funding entity *per se*.

A number of other elements are clearly central to the way in which Canada positions itself on the path forward. The consultations to date have identified the need for a common vision for an advanced Canadian DI, and have articulated three elements that should underpin this vision:

(1) Creation of a collaborative DI enterprise[2] that provides overall leadership and coordination;
(2) Development of a new approach to the stewardship and management of research data;
(3) Provision of a long-term funding and planning mechanisms that allow sustainable DI design, implementation, and operation.

---

[2] We use the term "enterprise" deliberately in the sense of a bold initiative of relatively grand scope, complexity, risk, and potential benefit, and which of necessity involves many people working collaboratively.

The following two sections lay out the first two phases of a draft action plan to address the specific DI challenges identified for Canada. It does so in a manner that respects the financial constraints on the research community and funders – e.g. as "the art of the possible."   Specific actions are proposed and roles and responsibilities of individual organizations are identified for discussion or to trigger other approaches. The Appendices lay the sort of timeline for action that could be considered, and a process for periodic review of progress against milestones.

Participants in DI Summit 2014 are asked to identify the key elements of an action plan that address the real problems of today, ideally also identifying how each of the stakeholders will contribute to delivery of the particular actions. The outcomes of the deliberations will inform follow up action by the Leadership Council following DI Summit 2014.

4. PHASE 1

4.1 Develop a collaborative national Coalition for going forward

An advanced DI will be difficult to achieve within the current fragmented environment. The most critical initial action by Summit participants is to explore what sort of collaborative mechanism would bring substantive concerted action and leadership to the DI enterprise – initially for a three year design or trial phase.

One approach is a Coalition that would build on the current DI Leadership Council, but with additional, or different, players brought into the mix in a more structured manner. While the Coalition would initially be guided by a preliminary DI Roadmap, a key responsibility would be to refine and expand upon any roadmap as a living document, to monitor and report to the community on effectiveness of action, and to assist with implementation of specific DI components as they come "on line".

The Coalition would provide – at a DI ecosystem level – high level liaison, coordination, planning, monitoring, reporting, communications, oversight and feedback to the community. The Coalition would not be a governing or management body, but one with "moral authority" and designed along the lines of collaborative Not for Profit Coalitions dedicated to achieving collective goals (See Appendix 1). It would contribute this moral authority to the individual and collaborative actions of the key agencies that have individual responsibilities related to delivery (as opposed to the primary funding sources) of the three main pillars of DI:

- Provision of high-speed network services and the related infrastructure.
  - o The obvious organizations to take the lead are CANARIE and the ORANS, with increased alignment among all network partners.
- Provision of HPC and related computing services (including implementation of new middleware and software tools).
  - o The obvious organization to lead is Compute Canada, but this will require attention to an improved funding model that provides a planning horizon and enhanced sustainability.

Advancing Canada's digital infrastructure //Améliorer l'infrastructure numérique du Canada

- Advanced research software and middleware development is a shared responsibility among CANARIE, Compute Canada, researchers funded by NSERC, SSHRC and CIHR and major national platforms funded by CFI under the MSI program.
- Provision of a policy framework and infrastructure for data management in all its aspects.
  - Key delivery and policy organizations are RDC, CKRN, CARL, and research institutions but there are numerous other organizations involved, and some significant structural challenges to overcome.
  - Within this, the meaning of "infrastructure" must be broadened (see below).

Both the DI enterprise overall, and each of three pillars individually, need to address the key principles found in the Policy Framework; i.e., leadership, coordination, funding, roles and responsibilities, and intended benefits.

The various national consultations (e.g. the Compute Canada and CANARIE strategic planning exercises) and initiatives (e.g. the CARL data management network and the work of RDC) currently underway would feed into the knowledge base and thinking of the Coalition. The Coalition would maintain an up to date inventory of such initiatives for easy access by policy makers, funders and researchers. While national in scope, the participating organizations and agencies would ensure regional and international engagement as appropriate and relevant.

The Leadership Council is currently a "coalition of the willing," without easy access to sufficient resources to deal with a national mandate and the complexity of DI issues facing Canada. A modest level of federal funding should be allocated to lever additional support from institutions and organizations for any coordinating body. This would enable a Coalition to function effectively, initially for the next three years in design mode as the various elements of the DI ecosystem are strengthened and move to a coordinated mode of action.

Roles and responsibilities relating to the Coalition:

Organization	Role
Industry Canada	Establish an overarching Policy Framework on data as a national asset and the principles for achieving a robust DI ecosystem as a foundational element of a national S&T and Innovation strategy. Work with TC3+ regarding resources for the Coalition.
TC3+	Establish guidelines and policies for DI, to ensure accountability and optimal access to, and use of, research findings supported by public funding. Consider providing resources (cost shared) for the Coalition during the DI design phase.
Leadership Council /	Develop a robust Coalition that represents the organizations

Coalition	responsible for implementation of the DI Roadmap. This would require:
	- New terms of reference - Review and readjustment of the membership - Identification of a neutral Chair (e.g. non-aligned with the key players) who can devote significant time to the work of the Coalition – ideally this individual will be a Champion of the DI, - Some support services (e.g., a permanent Secretariat during the DI design phase) - Clarity of initial roles and responsibilities for key players - Working groups to address key challenges.

4.2 Implement priority working groups

The DI leadership Coalition, whatever form it takes, is not the body to tackle all challenges related to DI coordination and implementation. It should be a nimble and efficient oversight body; its authorities should relate to coordination, monitoring and reporting and be broadly accepted; and its operational working groups should be constituted and re-constituted according to the immediate needs, not perpetuated as standing committees without priority purpose.

The substantive elements of its mandate would be addressed through expert *ad hoc* working groups formed for the tasks at hand, or the assignment of a specific mandate to existing groups. Those Working Groups would:

- Take action on gaps and issues within their existing mandates (e.g. alignment of strategic planning activities) and report on activities to the Coalition.
- Develop recommendations on specific issues identified by the Coalition and/or the funding agencies, reporting back to the Coalition for discussion, refinement, and approval of action plans.
- Implement those approved action plans, again reporting back to the Coalition.

An example of such working groups that engage a wide cross-section of the stakeholders is Research Data Canada (RDC), an organization constituted to address data management issues by bringing together organizations and researchers from academe, government, and the private sector into RDC working subcommittees. The constitution of the RDC reflects the fact that data are "everyone's responsibility"; its reach reflects the fact that there are many data-related activities that benefit from a forum for collaboration and sharing of best practice. The outcomes of the RDC deliberations, along with the strategic planning discussions of Compute Canada and CANARIE should be integrated into the Coalition agenda and DI Roadmap. In this context, the RDC would become the body representing the third pillar of the DI ecosystem, while not being a service delivery agent itself.

For any specific cross-cutting initiative, the Coalition would also identify the locus of implementation if it is not already obvious, acting as a coordinator of individuals and organizations that would not otherwise meet to discuss issues and options for solution. As critical early challenges are resolved, the Coalition will move on to dealing with other challenges and provide ongoing monitoring of effectiveness of action. The nature and membership of individual working groups will be tailored to fit the issue.

Roles and responsibilities:

Organization	Role
Leadership Council / Coalition	Identify key initial working groups, and develop their roles and responsibilities, reporting mechanisms, and accountabilities.
	Confirm the RDC role as a subcommittee to the Coalition for data management issues.
	Design mechanisms for approval and feedback to the Coalition and the broader community.
	Implement these governance activities.
	Create and dissolve working groups according to the need.

4.3 Pursue refinements to the DI funding system

The current project-based support for some individual components of DI and for overall DI leadership and development is unsustainable. Specific elements that need to be addressed include:

- *Provide long-range planning horizons*. Effective development of the DI ecosystem requires explicit commitment by the various funders to the importance of the three pillars of the national DI ecosystem (networks, computational capacity and data management, together with the "glue" of software/middleware), as well as the provision of a long-range planning horizon and long-range funding projections.

    o   In addition, such long-range planning (and funding) horizons will enable more effective coordination of planning among the pillars.
    o   Both CANARIE and Compute Canada require longer-term planning horizons and treatment as core elements of a national capacity in e-science (e.g., perhaps a rolling 10-year horizon with 5-year review points).

- *Develop explicit incentives for regional and national DI initiatives*. Incentives must be built into the funding programs to promote national and regional DI initiatives and to support pan-Canadian collaborative initiatives, even if sometimes at the expense of individual institutional thrusts. At present the incentives are unduly weighted towards individual projects, researchers, and institutions.

- This may imply a change to the nature of community consultations, RFAs, and Strategic Research Plans submitted by organizations to the CFI.
- The provinces will need to be engaged.
- It may also imply changes to how project-based research proposals are developed for the Tri-Councils, and/or coordinated with CFI and other major funding bodies such as Genome Canada.
- Five-year planning horizons are the minimum required.
- These collaborative efforts should be based on the value propositions discussed elsewhere in the DI documentation.

- *Expand the definition of infrastructure.* Recognize all of research data, data management tools, computing technology, software/middleware, networks, and skilled personnel supporting data management, training, and access to computational capacity as DI infrastructure.

  - Although there has been significant evolution in the CFI funding mechanisms in this direction, this may require further adjustment.
  - This expanded understanding of the nature of digital infrastructure should also inform the actions of Institutions, CUCCIO, CRKN, CARL and other key stakeholders.

- *Implement means to accelerate the development of research software and middleware.* Address the deficits in research software through a strategy for such software/middleware, better coordination of the existing providers of research software and middleware (including CANARIE, the ORANS, Compute Canada, Genome Canada, and others) and the provision of innovative funding programs that encourage the development of generic and domain specific tools that enable better return on existing investments in the technological infrastructure foundation.

- *Pursue innovative funding models* where warranted. While the TC3+ have been enormously effective at funding competition-driven support for science (and this must, of course, continue), the traditional competitive funding model is not an appropriate fit for the support of longer-term investments in system-wide initiatives in digital infrastructure. Some further innovation is required:

  - Predictability of opportunities for capital support is required, along with exploration of customized modes of support for Compute Canada as a national platform.
  - The current funding for capital acquisitions by Compute Canada is through project-focused competitions to the CFI, supported by co-funding through institutions and provinces. Renewal and operational funding is thus dependent on a good match with current CFI programming and the success of individual applications, neither of which may foster a balanced and effective evolution of

the national DI ecosystem (e.g., a recent proposal for Big Data in the medical and social sciences and humanities fields, and for confidential data, was unsuccessful).

- o The funding model for operating funding of national DI initiatives needs to be revisited to adjust matching funding requirements in the context of the scope and scale of activities, national S&T priorities and the interests of the partnering funders. While assurance of value is critical for investment decisions, for national initiatives, flexibility and negotiation may be in order to ensure national needs are met.

- *Introduce flexibility of allocation among elements of the DI ecosystem*. Provide a nimble and effective means to reallocate funding to those areas of the ecosystem that most need it, rather than keeping funding in silos. The needs will change over time, with evolution in technology and research frontiers. This entails at least:

  - o Means of adjustment to the balance of funding between acquisition and sustainable operations and maintenance (O&M), especially for regional and national initiatives.
  - o Means of rapid reallocation among the elements of the DI system: pipeline/networks, computing, data management, human resources, software and middleware.
  - o Consideration of infrastructure needs that are driven by long-range plans and priorities, as well as serendipitous opportunities that arise.

- *Recognize the costs of coordination*. Even relatively simple activities related to liaison, coordination, and collaboration can consume significant amounts of time and effort on the part of participating agencies and individuals who are going beyond their institutional mandates and acting as stewards of the national DI community. Not investing in such activities is a false economy.

  - o These in-kind costs must be recognized and addressed.
  - o Addressing the costs (including in-kind costs) of Champions is required; e.g., through provision of teaching/administrative leaves.

Roles and responsibilities:

Organization	Role
Industry Canada	Open discussions with the TC3, CFI, CANARIE and Compute Canada regarding how best to provide a long-range planning horizon and an effective, nimble funding environment for the DI ecosystem.
TC3+	Evolve the funding mechanisms to foster sustainability and overall coherence of the DI ecosystem in collaboration with IC and the Coalition.

4.4 Give priority to the data management pillar of the DI ecosystem

Of the three core DI pillars, data management is the one needing the most attention. Actions are required to address key gaps, including:

- *Develop data management strategies and policies.* There is a need for consistent, useful, "doable", and enforceable strategies and policies related to data storage, preservation, curation, sharing, use, re-use, and access.

    - These must reflect international trends and national/international interoperability regarding data standards, metadata and data citation for research data and research process data.
    - The full spectrum from "big" to "small" data need to be addressed, as each brings somewhat different challenges.
    - They need to address important issues related to privacy, confidentiality, proprietary IP, etc.

    The TC3+ consultation paper provides for an excellent framework incentive for fostering a culture of data stewardship for publicly funded research. The Coalition will need to monitor its implementation and barriers along the way.

    - Linkages with the Research Data Alliance (RDA) should be formalized.
    - CASRAI will play an important role in standards for data on the process of research data management.
    - The work of the RDC, in part informed by the outcomes of the RDA deliberations, will address many of the above policy gaps, but an implementation plan will need to be developed in conjunction with Institutions, scholarly and scientific societies and the TC3+.

- *Coordinate data management organizations.* Unlike the computational and network dimension of the DI ecosystem, the data management landscape is messier, with the involvement of many players:

    - The major policy body – the RDC – is not formally supported; its coordinating role needs modest support.
    - There is a growing body of opinion that much research data management infrastructure (including data repositories) will be distributed, but that institutions will need to be key players. This does not obviate the need for some national repositories to provide long-term and redundant storage and curation.
    - Some discipline and field-based data repositories are centralized (e.g. the Canadian Polar Data Network and the Canadian Astronomy Data Centre) but there are no formal linkages among such entities except recently through the RDC and its impressive engagement of stakeholders in its working groups.

- Librarians within research intensive institutions are actively engaged in taking on the role of data stewards, but face financial and human resource challenges in dealing with large-scale user access.
- Chief Information Officers are becoming more aware of the potential for them to take a more active role in data management, access, curation, preservation and storage in conjunction with their library colleagues.
- CRKN deals with digital text material but is looking to become an access point for diverse forms of digital resources including pre-publication research data.

There is therefore a need to develop a robust, accountable, and sustainable organization(s) with the appropriate agency and resources to design, develop, and eventually implement the data management component. This will require a champion to lead discussions among stakeholders to bring together the diverse constituencies in policy, delivery and funding structures that are better coordinated and more centralized than at present.

- *Share learning.* There are diverse, but often unconnected, experiments and innovations going on across the country dealing with different aspects of managing research data. Consider, for example, the BC and OCUL's cloud storage initiatives and the western university consortium (UofA, SFU and UBC) development of digital tools for preservation. The Coalition and the emergent coordinating organization for research data could play an important role in communicating successes and exploring how best to migrate good practices in a coordinated way across Canada.

- *Provide funding mechanisms to establish networks and centres of excellence in data management*. Creating a culture of stewardship for data as a critical national resource requires changes in how research is done. There is a need for centres of expertise and excellence focused on addressing challenges in digital curation that cannot be tackled by any single institution or discipline. Digital curation encompasses the active management of data throughout the research lifecycle. Such centres could also be the gateway to learning opportunities and other technical capacity as well as expertise on data curation. Networks and centres of excellence along the lines of the UK Data Curation Centre could be vital catalysts for the development of capacity, the emergence of a strong cadre of data scientists and the implementation of best practices. At present, however, data stewardship for all but the largest initiatives is mainly fragmented among researchers and their institutions, and support is therefore short-term.
  - The proposal in the TC3+ discussion document and the new initiative by CARL are excellent steps in this direction.
  - Provision of innovative funding mechanisms for investment in such initiatives will be required.

- *Overcome resource constraints.* With research data management emerging at a time when the overall funding system is under pressure, there are major constraints related to sustainable human and financial resources for the significant university investments required for training and implementing a transformation in capability.
  - Institutions will need to be active policy and financial players in ensuring a culture of data stewardship and in the delivery of effective approaches to data management.
  - The emergent coordinating body for research data management will require a modest level of support for a coordinated, functioning system.

Roles and responsibilities:

Organization	Role
Industry Canada	Within the framework of the federal S&T strategy (knowledge advantage), endorse the importance of research data management and the role of the TC3+ in implementation. As the major funder of the data creation, its preservation as a public asset is in the national interest. Explore with the TC3+ the role of the federal government in supporting research data management.
TC3+	Establish guidelines and policies for DI, to ensure accountability and optimal access to, and use of, research data supported by public funding. Develop appropriate funding policies (and, if necessary, programs and/or RFAs) at national, institutional, and project level related to all aspects of data management.
RDC (and its participants, including CRKN, CARL, CUCCIO, CASRAI)	Together explore how best to integrate the interests and activities of the diverse bodies active in the area and designate one coordinating centre or agency with appropriate authorities and resources. Develop outreach to the community to ensure clarity of its role and develop confidence in its authority.
RDC in conjunction with RDA and CASRAI	Develop data and metadata management standards, policies, and procedures for all users, making these consistent and transparent to the greatest extent possible across disciplines and sectors.
Discipline-based and other specialized bodies	Provide input to RDC/CARL/CUCCIO to refine policies and operational procedures related to their specialized interests. For example, develop: - Catalogues of important existing data repositories. - Inventories of existing financial and HR constraints in their areas. - Key issues related to their fields that must be addressed in

	global access and operations (e.g., proprietary, security, IP).

4.5 Articulate a value proposition

A clear, easily-understandable value proposition for a cohesive and sustainable digital infrastructure ecosystem will enable buy-in from academia as the primary initial DI users. This is necessary, but not sufficient: longer-term visions must be developed to engage partners and end-users in government and industry, as well as to foster socio-economic benefits for Canada.

Some specific attributes of the value proposition are:
- Must look "upstream" at the potential benefits of DI for advancing knowledge, but also "downstream" at potential impacts for the end-users of Big (and medium) Data in the government, industry, and general public domains, including open access to data.
- Research frontiers, associated data requirements, and end-user requirements should drive the technological DI investments, not the other way round.
- Emphasize the re-conceptualizing of "digital infrastructure" to include far more than hardware, and adjust value propositions, policies, and funding mechanisms accordingly:
  - DI is an interconnected web of networks, hardware, software, collaboration tools, data services and structures, skilled personnel and coordination bodies, and funding resources.

Development process - In developing the value proposition, best practices from other domains suggest that the process should:

- Engage users early on and actively in DI development, design, delivery, coordination and application (bridging any gaps between the designers and the users);
- Integrate the private sector as partners, users and funders;
- Develop common and more understandable language related to DI to make it more understandable and less intimidating for end-users and the larger stakeholder community who may not be as technologically knowledgeable; and
- Realistically assess the state of the DI environment, develop metrics for success (even if these are notional or "aspirational"), and assess the end-user communities' receptor capability (this will vary by discipline and sector).

Roles and responsibilities:

Organization	Role
Leadership Council / Coalition	Overall coordination and development. Value proposition for DI overall. Communication to funders.
CANARIE, ORANS, Compute Canada, RDC (or other data management	Value proposition for each individual element.

lead)	

5. PHASE 2

Tackling the phase 1 priorities will deliver a stronger, more coherent DI ecosystem than currently exists, addressing the majority of the high-level problems articulated in Section 2. Other issues that have been identified as requiring focused attention should be addressed as time and resources permit, some of which are described briefly below. The Summit participants are encouraged to expand upon these and identify other significant initiatives that should be incorporated in the forward action plan.

5.1 Engagement of government and private sector

Management of research data and access to networks and computational capacity are not problems restricted to the academic sector alone. Government research is facing similar challenges; government researchers and policy makers would benefit from improved access to, and sustainability of, research data generated outside its own walls. There are also efficiencies associated with increased access to shared resources, and benefits from the inter-sectoral and interdisciplinary connections that use of shared facilities often entails. Similar arguments hold true for the private sector. The recent CFI innovation of a Research Facilities search engine recognizes the importance of research collaborations that will help business innovate. This broader engagement in the DI agenda should be fostered by the Coalition/LC:

- *Private sector engagement:* Rather than being simply providers of infrastructure, the private sector is a potential user of the DI and partner in digitally-based and oriented research and commercialization initiatives. Through public-private partnerships, there is also the prospect of cost sharing, providing that the private sector clearly understands their potential returns on investment. There is a need for more formal recognition and promotion of the value of private sector engagement.

- *Government engagement:* As a research performer, a research user, a partner in research and a steward for the public interest, government (federal and provincial) has a vital interest in the quality and effectiveness of the DI ecosystem. A formal initiative should be launched to explore how best to ensure that the needs of provincial and federal governments can be effectively integrated into the ongoing discussions and priorities for evolution of the DI ecosystem.

5.2   Expertise and capacity development

- *Skilled expertise:* With the rapid introduction of digital technologies into all areas of research and scholarship, many researchers find themselves without the skills to access and utilize digital technologies, or to manage their data. Access to highly skilled support personnel, often with domain-specific training, is an increasingly important underpinning of the research ecosystem. The Coalition should consider working with

CARL, Compute Canada and CANARIE to explore how such services could be most efficiently delivered and coordinated.

- *Capacity development:* Training initiatives will pay big dividends for small investments, especially when focused on graduate students and early stage researchers. These training initiatives could be promoted by the Coalition and carried out under the auspices of CANARIE, Compute Canada, CARL and RDC, individually and collectively.

5.3   Middleware and software development

- Development of innovative middleware, advanced research software and other research tools is required. There is a need to document areas of unmet need and approaches to address them; identify areas of strength that could be grown, including with private sector partnerships; explore ways to expand awareness of sophisticated analysis tools among non-traditional users; and identify ways to promote experimentation and edginess.

5.4   Communications and engagement

- There is a need for more open and structured modes of engagement of DI leadership with academic research communities, scientific and scholarly societies, individual researchers, and research institutions, in the identification of gaps, priority needs and opportunities for the DI ecosystem.

6. AN ONGOING NEED

Within a foreseeable planning horizon, outputs from the DI enterprise are intended to be put into practice, in the field, by the Canadian research community. In the longer term, these outputs will be available to end-users in the private sector, governments at all levels, educators, and the general public. This means not only that the results are useful enough to encourage uptake by end-users, but also that mechanisms are put in place to foster knowledge and technology transfer and take-up (e.g., to deal with receptor capacity issues, fund intermediary steps such as proof of concept and demonstrations, address policy and privacy hurdles, etc. These aspects should be borne in mind within any of the initiatives discussed above, even if they are not explicitly mentioned.

The 2013-2014 consultation processes, Summits 2011-2012 and the various task forces and working groups active in DI will have brought together (virtually or in person) well over 300 influential opinion leaders across a variety of DI organizations, resources, and assets. This network asset should be nurtured, developed and consulted through regular communications as an appropriate national coalition is formed and the "go forward" DI Roadmap is developed. One result of the consultations has been an expectation within the sector that concrete actions will happen and DI-LC will deliver on the promise. Ongoing engagement with individuals within and external to this network will be critical to communicating and realizing the vision for the sector.

APPENDIX 1 – COALITIONS[3]

What is a coalition?
A coalition is an alliance of organizations, groups and associations working together to achieve a common purpose. There is a rich history of coalitions in the community and social welfare not-for-profit sector, but the concept has not been applied extensively in the post-secondary realm.

Coalitions may be formed as relatively short-lived alliances or longer-term partnerships. A coalition intended to have a longer lifespan will generally require more effort to maintain its organization and structure over time.

No matter how many issues it takes on or how long it plans to exist, the success of any coalition fundamentally depends upon the dedication of its members. Ultimately, coalition members must identify their common interests, articulate their shared goals, and work together to take advantage of the benefits that result from being part of a larger collective.

The value proposition for a coalition
- *Heightened effectiveness and community voice:* A well-organized coalition generally finds that is has more power and can exert more influence than an organization or individual operating alone.
- *Increased access to resources:* Working within a coalition provides opportunities for individuals and organizations to combine and share many types of valuable resources. Pooling resources allows coalition members to maximize their effectiveness by complementing one another.
- *Enhanced legitimacy:* Both within and outside of a community, the powerful presence of several different organizations banded together for a common purpose can attract more attention and command more respect than individual efforts.
- *Improved overall community organization and working relationships:* Coalition-building often creates opportunities for groups and individuals who had never worked together before to join forces and collaborate.

Success Factors
- Leadership style: Leaders with good knowledge base in the area, community respect, positive attitude, good negotiating skills, good communication skills, political sensitivity, the ability to guide the coalition toward the collaboration goal, while seeking to include and explore all points of view; belief in the process of collaboration, time.

---

[3] Adapted from *Strength in Numbers: A Guide to Building Community Coalitions.* Community Catalyst, Inc. 2003 http://www.communitycatalyst.org/doc-store/publications/strength_in_numbers_a_guide_to_building_community_coalitions_aug03.pdf

- Clear operational ground rules, e.g. for consensus development
- Unity of purpose and cohesion - a strong sense of solidarity and common purpose
- Collaborative spirit
- Effective communication within the collaboration and with stakeholders – both formal and informal
- Modest and cost-effective support tailored to the needs of the coalition
- Membership traits and participation - i) open-mindedness; ii) trust; iii) positive involvement; iv) personal commitment; and v) willingness to contribute to the success of the Coalition.

Notional TIMELINE and ACTIVITIES For DI Roadmap (an example of how the activities could be structured)

Legend:	Main actors	DI Roadmap Timelines
LC = Leadership Council (or augmented Coalition)		Notional roll-out of initiatives
TC3+ = Tri-Council Plus		Will be adjusted as a result of Summit discussions
IC = Industry Canada		
C = CANARIE & ORANS		
CC = Compute Canada		
D = Coordinated Data Organization (i.e., RDC/CARL+/CRKN/CASRAI etc)		
RDC = Research Data Canada		
GC = Genome Canada		
I = Institutions, with VPR and CIO leadership		
RDA = Research Data Alliance		

#	Activity	Main actors	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
			2014											**2015**											
1.0	**Develop national coalition**																								
1.1	Identify appropriate representation	LC, I, TC3+	X																						
1.2	Identify a neutral Chair/Champion	LC, I		X																					
1.3	Draft Terms of Reference (incl. roles and responsibilities)	All		X																					
1.4	Finalize ToR	All			X																				
1.5	Draft mechanisms for Coalition's operational mechanisms (e.g., "secretariat" functions, oversight, mandates and membership of working groups, liaison mechanisms, feedback, reporting, communications, feedback, etc.)	All		X																					
1.6	Finalize operational mechanisms	All			X	X																			
1.7	Endorse Policy Framework	LC, TC3+		X																					
1.8	Establish funding parameters for the Coalition	IC, TC3+		X	X	X																			
1.9	Interim coordination function in place	LC		X	X	X																			
1.1	Coalition coordination and oversign mechanisms implemented	LC					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2.0	**First Phase/Initial Working Groups (see WG details below)**																								
2.1	Identify/confirm priority challenges and associated mandates for working groups to address	LC		X	X																				
2.2	Create initial working groups (incl. ID of representatives, roles and responsibilities, accountabilities)	LC			X																				
2.3	Design mechanisms for approval of action plans, reporting to the LC/Coalition and broader community	LC			X																				
2.4	Launch working groups and communicate to wider research community with information on timelines and mechanisms of input	LC				X																			
2.5	Implement coordination and oversight (i.e., monitor WG activities, deliverables, etc.)	LC			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Advancing Canada's digital infrastructure //Améliorer l'infrastructure numérique du Canada

WG #1: Address major gaps & issues in data management pillar

#	Activity	Responsibility	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
3.0	*WG #1: Address major gaps & issues in data management pillar*																								
3.1	Assign WG#1 mandate to RDC, as formal subcommittee of the LC	LC			X																				
3.2	Confirm TOR, operational mechanisms (incl. Chair)	LC, WG#1/RDC			X																				
3.3	Develop mechanisms for coordination among organizations and institions; ensure broad acceptance	WG#1/RDC				X	X	X																	
3.4	Develop data management strategies and policies	WG#1/RDC, TC3+					X	X																	
3.5	Develop guidelines and policies to ensure accountability and optimal access to, and use of, research data	WG#1/RDC, TC3+					X	X																	
3.6	Task #1 -Assessment of existing Canadian initiatives, resonance with national strategy, scope and timelines for integration of best practices into national context	WG#1/RDC	X	X	X	X																			
3.7	Task #2 - Capitalizing on work of current RDC sub-committees, the current CARL Network initiative, and other Canadian inititiatives, ID residual gaps and initiate any additional sub-groups to address	WG#1/RDC	X	X	X	X	X																		
3.8	Task #3 - Work with instiututions to ensure effective means of implementing policies on digital scholarship - in particular - identifying critical barriers and approaches/incentives to address.	WG#1/RDC			X	X	X	X																	
3.9	Task #4 - Recommend options to the LC, institutions, the TC3+ and IC for effective support (functional and financial) of RDMI in the Canadian research environment	WG#1/RDC					X	X	X	X															
3.10	Task #5 - Ongoing. Position RDMI approaches in the international context; liaise with RDA, assess RDA outputs for Canada and ensure broad awareness of the deliberations and outcomes	WG#1/RDC, TC3+					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3.11	Task #6 Liaise with discipline- and field-based data repositories in Canada (e.g. polar data and astronomy) and internationally to ensure that best practices are integrated in Canada	WG#1	X	X	X	X	X	X																	
3.12	Task #8 Facilitate and promote the creation of initiatives that could become centres of excellence in one or more aspects of RDMI	WG#1									X	X	X	X	X	X									
3.13	Task #9 Review the arguments for and against the case for a national data repository that would serve as a redundant site and longer term storage/preservation/curation/access facility for critical national data assets	WG#1				X	X	X	X	X															
3.14	Other RDMI initiatives as appropriate over time. (There is much to be done on this front.)												X	X	X	X	X	X	X	X	X	X	X	X	X

#	Task	Owner							
4.0	**WG #2: Pursue structural and funding dimensions of DI funding system**								
4.1	Determine membership of WG#2	LC		X					
4.2	Confirm TOR, operational mechanisms (incl. Chair)	WG#2		X					
4.3	Task #1 - Examine and recommend an approach to providing effective planning horizons for DI funders, organizations and researchers while recognizing the requirements for accountability and flexibility	WG #2					X	X	
4.4	Task #2 - Examine & recommend how best to ensure a dynamic balance of funding among competing priorities, organizations, funders & initiatives (e.g. among institutional, provincial &d federal funders; between acquisition & operations/maintenance costs; among the three DI pillars; to respond to opportunities arising).	WG #2					X	X	
4.4	Task #3 - Recommend means of ensuring an appropriate balance between national, regional, institutional and project oriented initiatives, with an eye to achieving optimal return on DI investments. THis would include the articulation of effective incentives for collaborative and shared infrastructure.	WG #2					X	X	
4.5	Task #4 - While supporting the traditional approach of the TC3+ for competitive allocations, examine imporved models for capital and operating support of national DI infrastructure that provide a suitable planning and sustainability framework and do not pit project specific initiatives against the enabling capacity of generic infratsructure. Recommend changes to the IC and the TC3+. While effectiveness and efficiency must be reviewed, the decision on whether the function is desirable/required is a strategic one.	WG #2	X	X	X	X	X	X	
4.6	Task #5 - Review and recommend appropriate changes to the funding policies and programs of the TC3+ as a result of the evolution in the types of infrastructure required to support a data intensive research environment. A particular consideration is the mode of support for skilled non-faculty personnel.	WG #2	X	X	X	X	X	X	
4.9	Task #6 - Recommend improved mechanisms for promoting and funding the development of software and middleware tools, as well as means of ensuring their effective implementation and broad accessibility (including strategy, better coordination, innovative funding mechanisms)	WG #2			X	X	X	X	X
5.0	**WG #3: Develop value propositions**								
5.1	Confirm membership	LC		X					
5.2	Confirm TOR, operational mechanisms (incl. Chair)	WG #3		X					
5.3	Consult with broad community, individual disciplines, potential government and industry users	WG #3		X	X				
5.5	Draft value proposisions	WG #3				X			
5.6	Finalize value propositions	WG #3				X			
5.7	Draft performance metrics	WG #3				X			
5.8	Finalize performance metrics	WG #3				X			

Advancing Canada's digital infrastructure /Améliorer l'infrastructure numérique Canada

APPENDIX 3 – MONITORING PROGRESS

The draft DI Roadmap proposes an agenda that has a degree of urgency to it. Each year, discussions at annual CANARIE stakeholders meetings have reiterated the need for action and the observation that for progress there has to be a more integrated or "joined up" approach to the planning, funding, implementation and management of digital infrastructure among the various stakeholders. While this document does not advocate the creation of an additional operational or service delivery organization, it does stress the need for a collective or Coalition approach to developing the necessary cohesion and coordination. As this approach requires the commitment of some human and financial resources, albeit fairly modest, regular assessment of progress is required.

One role of the Coalition and the funder-observers should be to monitor progress along three axes:

- Timeliness and effectiveness of collaborative actions;
- Impact on the goal of a robust and sustainable DI ecosystem;
- Impact on the ultimate objective of a vibrant and effective science, technology and innovation system.

External feedback is also important. Representatives of the Coalition should participate in relevant community meetings at which there is an opportunity to discuss progress and continuing DI issues, including:

- The CANARIE annual Users' Forum
- Canada's Higher Education IT Conference (CANHEIT);
- The annual CUCCIO meeting;
- The annual CARL meeting;
- The annual CRKN meeting;
- The annual CASRAI Reconnect meeting;

APPENDIX 4 – THE COMPONENTS OF THE DIGITAL INFRASTRUCTURE ECOSYSTEM [4]

- **The framework conditions** – the policies and legal framework within which digital research is undertaken; the means of coordination and alignment of various components of the digital research environment; the suitability of funding systems and reward systems for pursuit of e-research; the capacity of Canada to deal with other international players in digital research.

- **Expertise and skills** (as an input) – the sufficiency and quality of skilled personnel, both generic and domain specific, for effective use of the components of the e-infrastructure.

- **Tools and services** – the software, applications and human support services that enable researchers to derive value from their data and to optimize the use of the research instruments and systems.

- **Research data and research data management infrastructure(RDMI)** – both data as infrastructure and systems for managing data—the collection, structuring, standardizing, archiving, curating and sharing of with system characteristics of flexibility, security, accessibility, interoperability, affordability, open access and high performance.

- **Computational facilities, tools and services** – hardware and associated software resources that enable both compute-intensive and data-intensive research, as well as the services and tools that enable value to be derived from the facilities. This includes both Cloud and Grid computing.

- **Networks and tools and services** – means of connecting researchers to data sources and transporting data among different locations.

- **Collaboration infrastructure and tools** – means of connecting researchers to researchers and partners in multi-sectoral research initiatives that are geographically dispersed and/or are utilizing common datasets and tools.

---

[4] From the October 16 2013 TriCouncil+ consultation document *Capitalizing on Big Data: Toward a Policy Framework for Advancing Digital Scholarship in Canada.*