Pre-Budget Submission to the Standing Committee on Finance

Increasing Canada's innovation performance by strengthening Canada's digital infrastructure ecosystem

Theme: Increasing the competitiveness of Canadian businesses through research, development, innovation and commercialization

canarie

Advancing Canada's Knowledge and Innovation Infrastructure

August 2014

Executive Summary

Canada stands at a unique moment in history as our long-term economic growth and competitiveness undergo a paradigm shift. In the digital economy, data has become the new natural resource and Canada must evolve its national digital infrastructure to effectively generate, store, and analyze vast amounts of data, transform it into knowledge, and use that knowledge to develop products and services that are globally competitive.

Canada's national ultra-high-speed network for research and innovation is a key pillar of digital infrastructure. The CANARIE network is the instrument that: connects and leverages other elements of Canada's digital infrastructure ecosystem; links the provinces; and connects Canada internationally. Along with its other complementary programs and services, CANARIE plays a critical role in enabling and accelerating knowledge creation, mobilization, and commercialization.

CANARIE's current funding is sunsetting in 2015. In order for Canada to continue to participate in data-intensive, collaborative, groundbreaking research and innovation, ongoing funding of CANARIE is required. Long-term funding of this valuable digital asset, which has been multi-purposed to support Canadians in both the private and public sectors, is an operational and a strategic investment in Canada's digital future.

While CANARIE's funding renewal is critical, we believe Canada should take this opportunity to further strengthen the overall digital infrastructure ecosystem by developing a strategy for data management and storage, and by promoting greater harmonization of the components in the ecosystem. The current funding approach to digital infrastructure has positioned Canada strongly, but a more strategic approach to digital infrastructure will ensure greater value from public investments, and fulfill the potential for economic and social benefits derived from these investments.

Introduction

A key driver of Canadian economic competitiveness is how well Canadian businesses embrace and leverage innovation to expand markets, increase competitiveness, and strengthen customer relationships. Innovation is most often driven by the creation and adoption of new knowledge. Today, powerful digital trends are converging and have a profound impact on how we create, disseminate, and preserve content and knowledge.

To capitalize on rapidly growing digital content requires the conversion of data to knowledge. World-class digital infrastructure, which overcomes the barriers of discipline, time, and location, is crucial for knowledge creation by enabling access to publicly and privately funded data.

Context

Successive Canadian governments have recognized the importance of investing in leading-edge research, infrastructure and talent to position Canada as a leader in new knowledge creation.

Science and Technology Excellence

Public science and technology investments, like the recently-announced $1.5B Canada First Research Excellence fund, amplify the need to ensure Canada's digital infrastructure is positioned to maximize the impact of these investments. To accelerate knowledge creation, knowledge mobilization, and greater commercial applications of new knowledge, there is an opportunity to harmonize the digital infrastructure that underpins research, and extend its impact to more users in the private sector.

Today's research is data-intensive and highly collaborative. Generally, the process follows four distinct steps:

- Data is generated. Whether information on water salinity from an ocean sensor, photographs from a satellite, or a collection of consumer data on buying habits, digital data is generated.
- Data is curated and stored. Currently, there is no central repository in Canada for data. Data is curated and stored within educational, government and private sector institutions, within departments in institutions, and by individuals.
- Data is "scrubbed" and analyzed. Data must often be cleaned or "scrubbed" to be made usable before algorithms are applied to manipulate it to respond to a specific query by a user.
- Data is visualized. The analyzed data is then presented so that it is usable, in the form of, for example, a graph, a model, or a video.

Digital infrastructure supporting this process is comprised of four components: advanced networking (to move data from its origin to its recipient, and to share data among researchers), high-performance computing (for advanced data analysis), data curation and storage repositories, and software applications that simplify the use of the other three components.

Overview of Current Digital Infrastructure Ecosystem

Network

CANARIE's core purpose is the advancement of Canada's knowledge and innovation infrastructure. The CANARIE network is the backbone of Canada's research and education network, a 23,000-km fibre optic backbone connecting 12 provincial and territorial networks and over 2,000 research and innovation institutions. CANARIE connects to Compute Canada's high performance computing facilities, and 100 international research and education networks, enabling Canadian participation in

world-class science and innovation, including groundbreaking initiatives in genomics, neurology, and high-energy particle physics.

Compute

Compute Canada, created in 2006 and funded through the Canada Foundation for Innovation, provides high performance computing facilities through four regional organizations across the country.

Data Curation and Storage

As yet, there is no pan-Canadian approach to data curation and storage, which limits opportunities for reuse of data. Data generated from research continues to be stranded across the country due to lack of defined and adopted practices for curation and storage, making much of these data unavailable for reuse.

Software

Software is critical, as it adapts data into human-understandable form, for example, through advanced data visualization tools. CANARIE has led the development of research software platforms and reusable software services to make access to data more intuitive and human-accessible. The vision is to further accelerate discovery, reduce duplication, and ensure research funds are applied to research, not to the development of supporting infrastructure.

Current Digital Infrastructure Ecosystem Challenges

While Canada's past investments in research and digital infrastructure have created a strong foundation, there are opportunities to improve the return on these investments. Increased harmonization among components of the digital infrastructure ecosystem would greatly improve system efficiencies. Consequences of the status quo include:

- Loss of knowledge – a recent report from the University of British Columbia concluded that 80% of scientific data is lost within two decades. Much of these data are unique to a time and place, and are irreplaceable, and many other datasets are expensive to regenerate[1]. As a result, new knowledge is not being fully leveraged and this contributes to a significant opportunity cost for Canada;
- Limited access to knowledge - lack of an integrated system with broadly adopted standards for data management and curation leads to inefficiencies and hampers cross-disciplinary collaboration. In addition, the absence of a national data "index" multiplies the risk of duplication of effort and the duplication of results;
- Constrained Canadian participation in global science – with respect to strategic digital infrastructure, Canada may be perceived as lagging behind

[1] http://news.ubc.ca/2013/12/19/scientific-data-lost-at-alarming-rate/

other jurisdictions (primarily the European Union). Canada may be a less attractive destination for leading researchers and top international students;
- Restricted strategic planning - the current approach to foundational digital infrastructure (compute and network) makes long-term infrastructure planning difficult;
- Missed opportunities for private sector engagement - Canada has an opportunity to incent private sector innovation by providing open access to new knowledge, and by strengthening opportunities for academic/private sector/government partnerships.

Strengthening Research and Innovation Leadership

Integrated Digital Infrastructure

Twenty years ago, one could consider the needs of computing, data, and networking as separate; today they are not. These three cornerstone technologies represent the DNA of modern digital infrastructure, and are so tightly intertwined as to be inseparable.

Canada has made leading-edge and strategic investments, which include CANARIE (1993), the Canada Foundation for Innovation (1998) and the Granting Councils, in support of research and innovation. Enhancing the interaction of these investments in support of an integrated digital infrastructure for basic and applied research excellence would realize significant benefits, accelerating knowledge creation, mobilization, and commercialization.

Following are some key issues and recommendations to be considered in evolving Canada's digital infrastructure ecosystem:

Issue: Ongoing funding and strategic alignment of foundational digital infrastructures

CANARIE and Compute Canada, as national digital infrastructure providers, enable Canada's most advanced, data-intensive, groundbreaking research. Both organizations are funded under different structures and operate under different mandates. While the organizations have worked to harmonize activities, there is an opportunity to more closely align the mandates, funding structures, and long-term objectives of Compute Canada and CANARIE. In its recent report, the Leadership Council for Digital Infrastructure strongly recommended the creation of a pan-Canadian plan for a sustainable and integrated system.

Recommendation:

CANARIE funding is sunsetting in 2015, and renewal of this investment is critical to Canada's digital infrastructure ecosystem, which is dependent upon the ultra high-speed network and programs that support cutting-edge knowledge creation and commercialization. CANARIE has a proud history of leading network innovations, championing research software development, and multi-purposing the network to enable private sector engagement in research and innovation. Ongoing investment in this critical asset is a strategic investment in Canada's digital future.

Further, Canada has an opportunity to strengthen and harmonize Canada's foundational digital infrastructures. This would mean an enhanced level of collaboration among research funders (Granting Councils), infrastructure funders (Industry Canada and Canada Foundation for Innovation), and other key players in the ecosystem. The result would be an efficient, effective and seamless national digital infrastructure supporting world-class, data-intensive research and innovation.

Issue: Limitations of the current approach to the data deluge

The creation of digital content is accelerating, much of it generated by government investments in research, including through the Granting Councils. Its increasingly pervasive use in research amplifies the current challenges of data management. There is a need to address current gaps in data management, curation and storage. As a principle, publicly funded research data needs to be FAIR: findable, accessible, interoperable and reusable, in order to achieve maximum utility.

Recommendation: The Open Science and Open Data initiatives are an impetus to evolve and implement a policy on access to research data. Canada has an opportunity to create a plan, developed in close consultation with the Granting Councils and other stakeholders, for research data curation, management and long-term storage. Along with a data management strategy, consideration should be given to the creation of a national system for data storage, so that research data may be shared using internationally-accepted data management standards and stored in reliable data repositories.

Issue: Limited private sector involvement in knowledge mobilization

In its latest report on the **State of Science, Technology and Innovation (2012)**, the Science, Technology, and Innovation Council (STIC) highlighted disappointing results in terms of traditional knowledge transfer indicators such as licensing and spinoff companies.

Providing access to research data to the private sector has emerged as one of the leading ways to incent greater knowledge transfer to the private sector and strengthen collaboration between academic, government and private sector partners.

Recommendation: Strengthen private sector engagement with the research community by reducing barriers in accessing research data, and increase the leverage of digital infrastructure by the private sector. One of the ways to facilitate access to data by all stakeholders, including the private sector, is to adopt data management standards so that data can be accessed and reused as easily as possible. This approach would open up opportunities to leverage municipal, provincial and federal data resources by both the academic and private sectors, as demonstrated by the Open Government initiative.

Conclusion

As the Federal Granting Councils declared in "Capitalizing on Big Data"[2], Canada stands "at a potential tipping point of a tremendous wave of exploration, innovation, productivity and growth, as individuals, post-secondary institutions, companies, governments and organizations of all types begin to exploit" emerging digital trends.

CANARIE and its partners in the digital infrastructure ecosystem are critical enablers of this wave. Ongoing investments in CANARIE, together with initiatives to enable Canadians to more easily access data and the digital infrastructures that transform data into knowledge, products and services, will strengthen the foundation supporting Canadian leadership in research and innovation.

[2] http://www.sshrc-crsh.gc.ca/about-au_sujet/publications/digital_scholarship_consultation_e.pdf