Data Innovation in Canada

Context for a National Conversation

Research and development is a key driver of Canada's economy, as it is for most countries, and data is a core ingredient in that effort. Research Data Canada works with multiple stakeholders to help ensure that Canadians have access to national and international collections of publicly funded research data, in support of innovation that benefits all Canadians. To maximize the social and economic impact of publicly funded research, RDC focuses on increasing stakeholder alignment, promoting support for data reusability, and addressing gaps in enabling technologies. It carries out this work in alignment with national efforts to advance Canada's digital research infrastructure as a whole.

This document is written on behalf of the RDC stakeholder community to facilitate the development of services to support researchers' use of data throughout the research lifecycle.

Data Innovation in Canada is intended to introduce a context for a national conversation involving stakeholders from across Canada's research ecosystem, and one that will help define a strategy for building our national infrastructure and services that intersect with provincial/ territorial interests and capacities. RDC will bring focus to this national vision through a series of consultations and conversations designed to help fill in the details over the coming months, and align RDC's stakeholder community around

EXAMPLES OF VALUE FROM ACCESS TO RESEARCH OUTPUTS

- **All Stakeholders:** immediate access to well-defined and documented data, enhancing or validating research, building knowledge
- **Researchers:** discovering previously hidden collaborations; finding useful/hidden data.
- **Funders:** building profiles of potential reviewers, maximizing funding dollars in active research areas.
- **Policymakers:** determining the impact of directed or general research funding, informing policy, evidence-based decision making.
- **Administrators:** quantifying and measuring organizational R&D outputs.

a common set of core principles and functions for building a national data management framework. RDC's primary locus is that the advent of digital technologies has had a substantial effect on the process of research, and has the potential to be truly transformative in how research is done, and thereby how new knowledge creation advances society and the place of the individual. RDC anchors this vision in a set of factors that must be addressed to support researchers across all domains in their efforts to discover, analyze, publish and reuse data.

Reusability

RDC will work with local, provincial, regional, national, and international organizations to promote accessibility and reusability. Reusability is the ultimate goal of robust and sustainable research data management services that drive an effective research ecosystem. To achieve this goal, we need to support reusability with policy frameworks that encourage an Open by Default approach to the dissemination of

research outputs, thereby encouraging greater innovation and economic development. This would include an investment in the funding of research data management and digital research infrastructure services and resources that support these efforts. A common suggestion of the amount needed to support this effort is 1.5-2% of total research funding. Given a recent review [1] that suggests that open data annually adds .5% of GDP more economic value than data that is closed or behind a fee wall, this investment would more than pay for itself.

Discovery

Discovery is the foundational pillar on which Canada's research ecosystem must be built, and the key to responding effectively to the reusability goal. Existing discovery systems typically function only at the institutional level, or in a few cases at a multi-institutional, provincial, or regional level. We need to ensure functionality from a national perspective. We also need to ensure that a national discovery layer, a system that provides access to all research outputs via a unified interface, is designed to answer questions from any user, whether a civil servant, university administrator, researcher, or interested civilian. Canadians should be able to easily access research evidence to find answers to their questions. This ability to discover research evidence is fundamental to advancing knowledge and innovation, and thereby Canadian society.

Metadata

A truly transformative approach to discovery would consider all research outputs throughout the research lifecycle, linked by common components and actors, including scholars, research resources, services, funders, and traditional outputs such as journal articles and data. By describing and harvesting the complete profile of research activity, we build a Metadata foundation that can leverage the power of new and emerging computing technologies to bring the smallest details of publicly funded research to the surface. If these metadata foundations are based on common standards and interoperable systems the challenges will be minimized. A key requirement of a system that stores this wide variety of outputs is to ensure that the security and privacy controls for both metadata and associated data is reflected in appropriate access controls and system-level functionality.

Preservation

For this to happen, one needs to ensure that the individual research outputs are preserved for the long-term, as the ultimate goal of discovery is to provide access to the actual data and information. An appropriate storage and preservation layer, or national Data Trust, is a necessary foundation of effective research data management. As with the analog collections that our national and institutional libraries have stewarded and preserved over centuries, we must recognize the need to invest in the same kind of long-term infrastructure for the digital realm.

---

[1] http://theodi.org/research-economic-value-open-paid-data

RDC believes that access to research outputs is just as critical as high-speed networks in terms of maintaining our leadership in research and innovation. We have become accustomed to supporting the long-term development and maintenance of the network; we need to bring the same philosophy to our storage infrastructures. If the preservation of research outputs alone does not justify the stewardship of this material for the long-term, then surely the opportunity to use this data to further research, discovery, and policy, will.

Coordination

Various organizations are engaged in conversations about developing data services at the national level. RDC provides opportunities for organizations spanning research agencies, universities and government to come together, to provide Coordination and develop a strategy for building the research data elements of Canada's digital research infrastructure, as well as to highlight the funds needed to deliver on that strategy. This coordinating effort also needs to be focused on developing a strategy for the long-term: one that clearly articulates that investing in the long-term, rather than project by project, is the only feasible way to proceed.

Finally, RDC is aware that with the majority of research funding coming from federal and provincial sources, and with the recent development of the Tri-Agency Policy on Open Access and their emerging policy on data management, there is some concern from the research community about the potential to increase the existing administrative burden and their institution's ability to support new requirements. Clearly the benefits outweigh these impacts, but without meaningful governance and due consideration to what changes may mean for researchers and their institutions, the implementation of transformative change will fall flat. Recognizing where funding can build momentum, and bring Canada in line with our international partners, is critical. A sustainable and Coordinated approach to transforming our national data fabric to support the Discovery and Reusability of well Preserved and interoperable research data and Metadata is key. RDC is committed to facilitating this transformation and working with our partners in Canada and abroad to achieve these goals.