National Consultation on Access to Scientific Research Data

Final Report

January 31, 2005

David F. Strong
and Peter B. Leach

NCASRD
National Consultation on
Access to Scientific
Research Data

Government of Canada Gouvernement du Canada

Canada

1 Executive Summary

Almost fifty years ago, the great writer and futurist H. G. Wells was very close to capturing both the development of the Internet and World Wide Web, and their potential impact on knowledge and research.

> *"Few people as yet, outside the world of expert librarians and museum curators and so forth, know how manageable well-ordered facts can be made, however multitudinous, and how swiftly and completely even the rarest visions and the most recondite matters can be recalled, once they have been put in place in a well-ordered scheme of reference and reproduction."*
>
> — H. G. Wells
> "*World Brain: The Idea of a Permanent World Encyclopaedia*"
> Contribution to the new *Encyclopédie Française*
> August 1937

That day has arrived, and Canada must seize it.

In mid-June, an expert Task Force, appointed by the National Research Council Canada (NRC), came together in Ottawa to plan a national Forum as the focus of the National Consultation on Access to Scientific Research Data (NCASRD). The Forum brought together more than seventy leaders Canada-wide in research, data management, administration, intellectual property and other important areas.

This Report is a comprehensive review of the issues, opportunities and challenges identified at the Forum, complemented by a selection of the supporting documents presented as Appendices.

1.1 The New World

Complex and rich arrays of scientific databases are changing how research is done, speeding discovery and creating new concepts. Increased access will accelerate these changes, creating a new world of research **and a whole new world**. When these databases are combined within and between disciplines and countries, fundamental leaps in knowledge can occur that transform our understanding of life, the world and the universe.

For example, in the analysis of human genetics, the technology to capture enormous amounts of data and to mine them for new information is already showing the genetic make-up of life and the understanding of numerous diseases and syndromes. We will soon be able to analyze such complexities as the pre-disposition to disease in animal and plant populations based on genetics, social and environmental conditions, and demographics, so that all these factors can become part of new disease prevention strategies. With the ability to access and integrate data compiled in different fields, totally new knowledge regimes are being opened in ways that have historically been impossible.

1.2 Canada

For Canada to be a leader in the knowledge economy, the country must be a leader in the new world of research. For Canada to lead in this research transformation, it is essential to take swift action on the recommendations of this Report. For Canada to benefit economically and socially, substantial changes are required in our scientific enterprise, including:

- research culture and behaviour;
- research institute management, policies and strategies;
- legal and policy frameworks;
- financing and budgeting of research; and
- data technologies and computing infrastructure.

Some of our OECD competitors are moving on these challenges much faster, posing an ultimate and very real threat to Canada's economic and social well-being.

While Canada is involved in many global database and research initiatives, these are individual cases not bound by any national strategy or standards. As a result, much of the data on which our knowledge is being built today is hard to access by other Canadian research communities, and is often not ideally structured to be as useful or as open as possible, even within the discipline for which it is being constructed. The vanguard of Canada's national activity and international presence in access to scientific data is through the shared leadership of the Canadian National Committee for CODATA (CNC/CODATA), the Canada Institute for Scientific and Technical Information (CISTI) and the Canadian Association of Research Libraries (CARL).

Member institutions of CARL are already active in preserving some of the country's scientific heritage in digital format, and provide much of the knowledge that is being used in the data capture, access and preservation processes of many national and international scientific database projects. CARL members are 27 of Canada's major academic research libraries, together with CISTI, Library and Archives Canada (LAC) and the Library of Parliament.

1.3 The Past

However, no national data preservation organization exists, nor does Canada have any national data access strategy or policies. Participants at the NCASRD expressed considerable concern about the loss of data, both as national assets and definitive longitudinal baselines for the measurement of changes over time. These losses occur as a result of storage media degradation, media and metadata loss, and software and hardware obsolescence, as well as privacy policies and decisions (e.g., by research ethics committees), and a lack of planning or attention to preservation beyond the individual researcher or organization.

1.4 Action

So, action is urgently needed to stop such degradation and loss of the country's research heritage; action that could concomitantly thrust Canada into a leading position in this new paradigm for research and development.

We recommend the creation of a task force, dubbed **Data Force**, to prepare a full national implementation strategy, and mount a pilot project to show the value and impact of multi-person and multidisciplinary access to research data. Once such a national strategy is broadly supported and has obtained appropriate funding commitments, we propose the establishment of a dedicated national infrastructure, tentatively called **Data Canada**, to assume overall leadership in the development and execution of a strategic plan. The plan would encompass and presumably extend the NCASRD's recommendations.

1.5 The Future

With Data Canada implementing the recommendations of this Report, we believe the country will be able to achieve the NCASRD's Vision of Canada's place in the global research enterprise of 2020. We envision that by then:

Canada is the centre of a global knowledge grid. It has become the desired nation with which to partner in research, because of its national system of open access to research data. Through this system and the collaborative culture it has generated, Canadian creativity and innovation are best in class worldwide. Open, but secure, access to powerful and globally assembled data has transformed scientific research. Researchers routinely analyze problems of previously unimaginable complexity in months, rather than decades, leading to revelations of knowledge and discovery that have enriched quality of life, transformed healthcare, improved social equality, provided greater security, broadened decision perspectives for social, environmental, and economic policy and advancement, and transformed the advancement of human knowledge.

Canada is not alone in having such lofty aspirations and delays in action will cause cumulative damage to its potential leadership role. With other countries already taking progressive action, delay is a destructive option.