

Towards a Robust and Sustainable Ecosystem for Enhanced Digital Scholarship: From Common Understanding to Action Plan

Chad Gaffield, President, SSHRC

October 10, 2012

I) Introduction:

Informed by recent conferences, reports and studies, this paper offers a description of the key features of a robust and sustainable ecosystem for enhanced Digital Scholarship in Canada. This description reflects the fact that digital technologies have increasingly enabled, accelerated and influenced a new integrated model of scholarship that transcends the 20th century distinction between teaching and research, and between campus and community. Students, professors, research partners and those in the larger society are now being connected formally and informally in efforts to learn about all aspects of the past and present, and to use such learning to help make a better future. In this context, `research data` are now also `learning data` as well as `innovation data.` This major change which can perhaps be best captured by the expression, `Digital Scholarship,` must be taken into account in the policies and practices at both the federal research granting agencies and post-secondary institutions (many of which were developed within an analog paradigm).

For this purpose, Digital Scholarship is defined inclusively as the digital technologies, digital content, and digital literacies used in education, research and innovation involving the post-secondary sector. The paper then proposes an Action Plan with specific responsibilities for federal government granting agencies and post-secondary institutions.

In the first instance, the objective of this paper is to help an emerging `coalition of the like-minded` (especially those in positions of authority and with the responsibility and accountability to take action) agree on a common understanding of the key features of the desired ecosystem. In the second instance, the objective is to help the emerging `coalition of the like-minded` move from a common understanding to an Action Plan that includes commitments to take specific steps forward each year over the next three years in a coordinated way.

Towards these two objectives, this paper is a point of departure for composing a collectively-authored working document by the following self-selected members of a coalition of the like-minded: the Presidents of the relevant federal research funding agencies (SSHRC, NSERC, and CIHR together with CFI, CANARIE, Genome Canada, NRC-CISTI, etc); post-secondary institutional leaders (CIOs, Librarians-Archivists, VP-Rs, and Provosts); and other contributors to digital scholarship (CASRAI, etc. as well as those across the health field such as Academic Health Sciences Centres, community-based health care settings, CIHI, etc.). Once this coalition agrees on the content, the document could be opened in a wiki for further enhancement by other interested parties including those working at the provincial level and in the private sector.

The overall approach builds upon two successful initiatives of enduring value for digital scholarship that were taken by federal agencies and post-secondary institutions in Canada in the 1990s: the Data Liberation Initiative and the Canadian Research Data Center Network. The proposed approach requires coordination at the institutional level and among federal funders in

the implementation of an agreed-upon overall strategy that will be continuously updated in keeping with the dynamic character of Digital Scholarship.

II) Background:

Taken together, the multiple conferences, reports and studies of recent years suggest the emergence of a common understanding about the key features of a robust and sustainable ecosystem for digital scholarship in Canada. While only in the nascent stages, this ecosystem already transcends campuses just as it crosses geo-political borders in the increasingly networked and 'open' world of international efforts to advance knowledge and build understanding about all aspects of the past and present with a view toward creating a better future.

At the same time, the meetings and reports of recent years point to the essential roles played in Canada by two entities in supporting the development of digital scholarship: post-secondary institutions (including Academic Health Sciences Centres) and the federal research funding agencies (SSHRC, NSERC, and CIHR together with CFI, CANARIE, Genome Canada, NRC-CISTI, etc). These two entities have funded thousands of research initiatives as well as bodies such as Compute Canada Calcul (and the regional components), the Canadian Association of Research Libraries, CUCCIO, the meetings of VP-Rs and Provosts, the Canadian Research Data Centre Network, the Research Data Strategy Working Group, CASRAI, and other key contributors to digital scholarship. Moreover, federal research funding agencies and post-secondary institutions have collaborated with provincial bodies as well as the private sector to advance digital scholarship.

In playing their leading roles, the federal research funding agencies and the post-secondary institutions make different contributions based on their distinctive mandates. Federal funding agencies 1) take a pan-Canadian perspective on funding decisions; 2) they provide support for specific initiatives to advance knowledge, develop talent and enhance innovation during specified time periods; and 3) they work within the larger context of the federal government.

Post-secondary institutions 1) take an institutional perspective on digital scholarship with special connections to their host region within a national and international context; 2) provide support for both short-term and enduring initiatives to advance knowledge, develop talent, and enhance innovation through digital scholarship; and 3) operate primarily within a provincial or multi-provincial context.

In other words, ensuring pan-Canadian and internationally-connected digital scholarship is primarily a preoccupation of the federal research funding granting agencies. This is why they fund bodies such as Compute Canada Calcul and CASRAI. In contrast, ensuring sustainable digital scholarship (as nodes within regional, national and international networks) is primarily the responsibility of post-secondary institutions (in keeping with the traditional role played by campus libraries, laboratories, research centres and institutes).

With a view toward effective coordination, the proposed approach is to delineate clearly the responsibilities for specific features of the desired ecosystem based on a common understanding of the overall work to be done. The central question to be addressed is: what are the specific

features of a robust and sustainable eco-system for Digital Scholarship for Canada, and who is best situated to assume responsibility for each one?

Together, both federal research funding agencies and post-secondary institutions also advance digital scholarship through specific collaborative initiatives for certain time periods (characteristically chosen through funding opportunities involving merit review by agencies of proposals submitted by institutions). These initiatives are characteristically funded for no more than seven years.

Moreover, post-secondary institutions and federal research granting agencies collaborate and share jurisdiction on certain major research infrastructures and platforms. And, in certain cases, federal investments have, and can be expected to, extend beyond the characteristic limit of seven years.

Perhaps most importantly, post-secondary institutions have a mandate to protect the digital assets with which they have invested or been entrusted. This is their digital stewardship responsibility. In a complementary way, federal granting agencies have a mandate to ensure that digital assets produced through public funds are deposited with institutions that provide digital stewardship.

In this context, the urgent question has now become: how can the relevant federal research funding agencies and post-secondary institutions collaborate most effectively in building a robust and sustainable ecosystem for digital scholarship? How can the bodies funded by these two entities (Compute Canada Calcul, CASRAI, CARL, CUCCIO, etc) best help them fulfil their complementary mandates? How can effective partnerships with the provinces and the private sector be advanced as part of this collaboration?

III) Towards a Common Understanding of the Key Features of a Robust and Sustainable Eco-System for Digital Scholarship: Key observations from recent conferences, reports and studies:

- 1) Digital scholarship is now evident across all fields of enquiry as data become the 'coin of the 21st century realm.' From studies of colliding particles to research on human thought and behaviour, the importance of digital scholarship is increasing rapidly across campus and beyond.
- 2) Along the way, data are now understood in terms of numbers, words, images, sounds, and, indeed, digital representations of all human and non-human phenomena. In this context, the future will require increasingly sophisticated approaches/mechanisms for interrelating and integrating multiple datasets; for instance, researchers are now focusing on integrating clinical/health data with diverse environmental, socioeconomic or genomics data.

- 3) The use, re-use and re-purposing of data are becoming increasingly important articulations of the deep cultural changes now underway in education, research and innovation.
- 4) As distinctions between creators and users blur, data often exist in a dynamic rather than fixed state as a result of multiple and iterative engagement. These changes call for updated and new policies and practices both in post-secondary institutions and federal research granting agencies. As illustrated by health research, complex privacy, confidentiality and ethical barriers arise around access to and use of personal and collective data across multiple researchers/research settings. Since many legal and ethical considerations across diverse research areas are shared internationally, it is especially important that Canada engage with partners around the world in order to develop appropriate policies and practices for digital scholarship.
- 5) The focus of attention on digital scholarship has moved from technology to an interrelated technology, content, and user-oriented nexus. Most observers conclude that digital technology is far more advanced at the moment than either the availability of digital content or the support for users (including digital literacy education), and that therefore considerable attention must now be given to correct this imbalance. While continued improvement in computing processing and connectivity is essential especially given the increasing importance of massive data, special attention needs to be paid to enhancing access and use both on campus and beyond. Perhaps the most urgent need is for skilled and sophisticated people who can work effectively in a digital environment including technologies, content and literacies (from access to analytics). An integrated approach to the connectedness of technology, content and use in digital scholarship will accelerate its transformative potential to enhance life in Canada and beyond.
- 6) While discussion in the 1970s-1990s often stressed the importance of standardization in a 19th and 20th century cookie-cutter sense, the new emphasis is on both-and solutions to the problems of preservation, interoperability, meta-data, data delivery systems, user interfaces, etc. In the distributed, open, empowered world of digital scholarship, coordination rather than control is key especially as networks replace vertically-integrated hierarchies. Such coordination is not only essential domestically but also internationally since scholarship now transcends geo-political borders to an extent that goes well-beyond connections of the 19th and 20th centuries.
- 7) Great unevenness characterizes the digital scholarship landscape especially with respect to preservation infrastructure which has become the most urgent need to support and sustain digital scholarship. The most intensely cultivated are fields in health where certain domain repositories, such as bioinformatics, systems biology, structural biology, and large genomics, including human genome, and disease

datasets are already robust, well-validated, well-curated, and kept up to date by specified organizations based on researchers inputting data according to well defined standards. Similarly, domain repositories in astronomy and crystallography are well-developed in the natural sciences as are certain topical repositories such as the Canadian Polar Data Network. Despite the intense recent work in fields such as Digital Humanities, much more work needs to be done with respect to data reflecting human thought and behaviour in order to build on the early success of the Data Liberation Initiative and the Canadian Research Data Centers.

- 8) In ways similar to the changes in the recorded music industry, digital scholarship is moving from an emphasis on data ownership to a provision of data services. Whereas libraries once held and lent scholarly journals, for example, they now increasingly provide access to digital publications. However, questions of ownership, curation, and access as well as business models for sustaining a `services` approach (that includes support for analytics, visualization, etc) remain far from resolved. Current policies and practices developed for analog scholarship (including funding possibilities) leave major gaps as well as unnecessary duplication for digital scholarship.
- 9) While the increased connectivity of digital scholarship has in many ways made physical location less relevant, the growing importance of massive data which cannot be moved easily has increased the importance of location. In addition, the growing recognition of the importance of face-to-face interaction as a component of effective computer-enhanced learning and collaboration helps explain the importance of regional clusters with networked institutions that are then linked to remote locations nationally and internationally.

IV) Action Plan for Harvesting Low-Hanging-Fruit (Fall 2012 to Spring 2013):

- i) Federal Research Granting Agencies:
 - 1) Survey how other like-minded countries are dealing with Digital Scholarship with a view toward learning from their strategies and experience (including nature and extent of investments), and in order to bench mark progress and to monitor how Canada is doing relative to others.
 - 2) Informed by precedents elsewhere, publish guidelines for effective Data Management Plans to foster a culture of scholarly social responsibility, and make compulsory their inclusion with grant applications and final reports.
 - 3) Informed by precedents elsewhere, formally describe contributions to digital scholarship that will be included in the evaluation process (creation and deposit of data in trusted repository, etc.).
 - 4) Consider funding a small number of centres of excellence in Digital Scholarship to advance policies and practices (along the lines of the Knowledge Impact in Society program that cultivated innovation in campus-community connections by funding 6 centers across Canada).

- 5) Convene annual meetings with provincial counterparts to focus on planning and policy.
- 6) Engage with private sector leaders to realize more fully the `innovation data` potential of Digital Scholarship.

Post-secondary institutions:

- 1) Identify institutional and trans-institutional gaps that need to be addressed in order to ensure a sustainable and comprehensive preservation infrastructure for digital scholarship;
- 2) Establish a list of post-secondary trusted data repositories that have received the seal of approval; commit to milestones for achieving full coverage across Canada by 2014;
- 3) Informed by precedents elsewhere, identify major gaps in the treatment of legal, ethical, privacy and confidentiality issues in Digital Scholarship;
- 4) Contribute to the building of inventories of scholarly digital resources using citation fields based on emerging national data standards;
- 5) Update undergraduate and graduate programs to enhance digital literacy including deep analytic expertise, data management, and the operation of research computing systems.

Collaborative activities between Federal Research Granting Agencies and Post-secondary institutions:

- 1) Review and update how the bodies funded by them (Compute Canada Calcul, CASRAI, CARL, CUCCIO, etc) can best fulfil their complementary mandates and identify gaps that require attention (e.g. international linkages).
- 2) Work together to roll out the requirement for Data Management plans, including the identification of services that post-secondary institutions would provide to support Data Management plans; work together to establish with CASRAI a metadata standard describing the elements of a Data Management plan; and work together to develop tools for preparing Data Management plans.
- 3) Work together to prepare criteria for selecting data for preservation.
- 4) Work together to identify certified repositories that Granting Agencies will reference.
- 5) Work together to develop solutions for the treatment of legal, ethical, privacy and confidentiality issues in Digital Scholarship.
- 6) Establish a biennial forum to review the agreed-upon common agenda and work plan, and to report on steps forward.