CSDH/SCHN Submission to
Industry Canada Consultation:
Developing a Digital Research
Infrastructure Strategy

The Canadian Society for Digital Humanities / Société canadienne des humanités numériques (CSDH/SCHN) thanks Industry Canada for a chance to provide a submission to the consultation on Developing a Digital Research Infrastructure Strategy. CSDH/SCHN represents researchers across the humanities and across Canadian universities who are using digital methods in the study of the cultural record. We study everything from current social media to Canadian literature. DRI is important to our capacity to do research and to train students in the use of digital methods. We have organized our submission around your questions:

- How can DRI be realistically transformed, strengthened and supported over the next five years?

  Crucial to a DRI Strategy that can position us as a "big data" leader is the capacity to manage and enrich the data that matters to Canadians from health data to our stories and histories. We encourage Industry Canada to therefore work with libraries, archives, and other holders of information about Canadians to develop a Strategy that encourages the coordinated stewardship of research data. This means involving those that understand the enrichment and documentation of the data. It means understanding the ethical and privacy issues around data. It means establishing a network of Curation Centres so that all researchers, both community and academic can deposit research data.

- What are the biggest challenges limiting the effectiveness of the DRI ecosystem? What opportunities are there to more efficiently deploy the human, technical and financial resources currently being devoted to DRI? How, and in what priority, should they be addressed?

  The biggest challenges we see include A) The availability of good training, advice, and support is uneven. B) There is a disconnect between organizations like Compute Canada provide research processing and those like library depositories that provide the long-term preservation. The establishment of a network of regional Data Curation Centres developed with research libraries could be coordinated with Compute Canada to provide integration. A researcher should have an integrated environment that lets them move data from the infrastructure used to gather and study to that used to preserve as needed. Such a network of regional centres would have to be funded through coordinated national and provincial support so we don't end up with uneven research infrastructure.

- What do you see as the biggest challenges to effective data management and the development of data standards in Canada? What could be done to promote a more rigorous and coordinated data management system that supports research excellence and maximizes the benefits generated by our investments?

  The biggest challenge is lack of engagement among researchers. Many researchers, if they have heard of data curation at all, see it as one more task that they can't afford to do without support. A second challenge is the uneven availability of advice, training, and support. A way to support excellence would be to reward it with excellent support from the moment a research project is imagined to when it is finished. This would mean having good advising available to all researchers. It would also mean funding the extra work of properly documenting and depositing research data. We recommend that all the research councils have small Archiving Grants available to funded

research projects to collaborate with their libraries in properly curating their data.

Another longer-term solution is to weave training in digital curation into the disciplines that deal with information, especially the social sciences, arts, and humanities where the cultural/literary/historical record is studied and enriched. IC could provide support to the granting agencies to support graduate student training across the disciplines.

- What is the current capacity within post-secondary institutions to support research data curation?

The current capacity varies widely. Certain large research-intensive universities have excellent training, support, and local infrastructure. Others have little. We recommend IC support the formation of a consortium of regional centres of excellence (Data Curation Centres) that can provide training and support. These would be built on existing centres of strength and could be coordinated with Compute Canada.

More importantly, there is little understanding among researchers in the humanities of the need for stewardship. Many researchers still think of print publications as the goal of research and hoard their data. Associations like CSDH/SCHN can play a role in changing the culture among the academics who do the research that gathers the data. Training the next generation of reseachers is crucial. Compute Canada is starting a Software Carpentry national training series; we need a Data Carpentry series to mirror it.

- What are the biggest strengths of the DRI ecosystem? How will these strengths be affected and prioritized by a transformation of DRI in Canada?

Canada is recognized for its strength in humanities computing with important centres and projects across the country. We have a particular strength in the development of tools for both scholarly editing of textual data

and for the analysis of textual data. We have a tradition of training graduate students through programs ranging from the MA in Humanities Computing at the U of Alberta to the Digital Humanities Summer Institute at the U of Victoria. Flexible support for the training of HQP could build on Canadian successes. A DRI Strategy should be careful to not focus on only one form of training and support, but support a rich ecosystem.

- What is the role of the private sector in supporting a strong DRI ecosystem in Canada?

  The private sector has played and can continue to play a role in the development of innovative tools, especially when they provide tools to researchers in a form that can be adapted to research. Companies that invested in analytical tools have partnered with academic networks with CFI support – these partnerships should continue and be fostered. The private sector has also supported internships through programmes like MITACS and should be encouraged to continue to do so. For Canada to be a big data leader we need to connect students to industry to the benefit of all.

Do you have any other comments or suggestions to support the development of the DRI strategy?

  Data about Canadians, our governments, our histories, our stories, and our ideas is a public good that should be as accessible as possible to all. A DRI Strategy should not just focus on preserving data from academic researchers for academic researchers. A DRI Strategy should make data accessible to all Canadians. Imagine a model where any Canadian could deposit the stories, pictures, or data that mattered to them. Imagine a model where any Canadian could ask questions of their governmental data, the historical data about their community, and the stories of their communities, where appropriate.

  Article 19 of the Universal Declaration of Human Rights states, "Everyone has the right to freedom of opinion and expression; this right includes freedom to

hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers." Let's create an Open Canada model that is inclusive and free. Then we will be a leader.

September 2015