

2016

# The Academic Data Librarian Profession in Canada: History and Future Directions

S Vincent Gray

*Western University*, vince@uwo.ca

Elizabeth Hill

*Western University*, ethill@uwo.ca

Follow this and additional works at: <http://ir.lib.uwo.ca/wlpub>



Part of the [Archival Science Commons](#)

---

## Citation of this paper:

Gray, S Vincent and Hill, Elizabeth, "The Academic Data Librarian Profession in Canada: History and Future Directions" (2016).

*Western Libraries Publications*. Paper 49.

<http://ir.lib.uwo.ca/wlpub/49>

# The Academic Data Librarian Profession in Canada:

## History and Future Directions

*Elizabeth Hill and Vincent Gray*

FROM THE 1970S onward, Canadians have been active in developing services and establishing structures to support the dissemination of data. In recent years the academic data profession in Canada has largely developed around access to data from the national statistics agency, Statistics Canada, and around the services which have been developed to permit access to these data.<sup>†</sup> This chapter will provide a historical background for these activities and explain how current and emerging trends continue to affect the profession.

Data librarianship in Canadian universities in the 1970s and 1980s was influenced by factors similar to those in other countries. As was happening elsewhere, researchers, particularly in the social sciences, were beginning to collect and analyze machine-readable (MR) data files (MRDF)<sup>‡</sup> and to use already-collected data files as cost-effective resources for secondary analysis. Academic researchers, especially in the social sciences, increasingly looked to Statistics Canada as a source of MRDF for analysis.

### The Development of Data Services

Initially, data services were typically located within computing centres as opposed to libraries, and librarians were seldom involved in the delivery of data. Institu-

---

<sup>†</sup> Western University (formerly University of Western Ontario) has been a participant in and/or developer of these services both prior and during the tenure of the authors, who have combined for over 45 years of data experience at Western.

<sup>‡</sup> Results of surveys or censuses coded into machine-readable format such as the Social Change in Canada Survey, 1977, conducted by Institute for Behavioral Research at York University.

tions typically chose to establish data services in pre-existing computing centres due to the cost and complexity of computers. The University of British Columbia established a Statistical Centre in 1963–1964.<sup>1</sup> Carleton University's data service, which in 2015 celebrated its 50th anniversary, was the only Canadian record in the publication *Social Science Data Archives in the United States, 1967*. Western University's data service was founded in 1972, in the Social Science Computing Laboratory (SSCL).

An article in a 1982 *Library Trends* special issue on data records that discussions began in the 1960s on how libraries or social science data archives might support data-related inquiries and provide services to support researchers.<sup>2</sup> Initially, discussions occurred within archives as libraries were not seen as logical partners, lacking the technology necessary to support statistical software programs. A second article in the same issue of *Library Trends* reported attempts to influence academic libraries to take responsibility for collections and services pertaining to data.<sup>3</sup>

The 1972 merger of the University of British Columbia's computing centre's data service into the university's library created the first Canadian library-based data service.<sup>4</sup> Generally, however, Canadian university libraries did not see data as part of their service mandate until the mid-1980s. In a 1979 article, Slavko Manojlovich recognized the growing use of computers for data analysis and that machine-readable files from the census were becoming increasingly important to researchers.<sup>5</sup> Canadian census data had been first released in machine-readable format with user summary tapes and public use sample tapes for the 1961 census. Manojlovich contended that the library, being the major depository of information at universities, was the logical home for MRDFs. Using a survey of Canadian government documents librarians, he found that 35 Canadian universities had census data files. Services to assist researchers with the data were found in disparate locations from university to university. Overall he recognized the need to have an information professional mediate access to MR data.

## Locating and Providing Access to Data

The model for accessing and acquiring data during these early years was very different from today. If a researcher realized that Statistics Canada had a relevant data file, he or she might purchase it (usually for Can\$300 or less), or contact Statistics Canada staff who could, and often would, provide the data for free. Other Canadian providers included the Machine-Readable Archives (subsequently the Government Archives) Division at the National Archives, and specialized centres, such as the Health and Leisure Databank at Waterloo and survey institutes at York, Alberta and Winnipeg, which would produce print catalogues of their holdings.

Researchers usually ordered data independent of any university-wide collections policy or service point; generally, they and their graduate students would be the only users of these files, so files might be duplicated on campus when acquired by different research teams. Data files would arrive on magnetic tapes, which required computer mainframe-based tape drives to read, together with technical staff and the infrastructure to support access and analysis. The Inter-university Consortium for Political and Social Research (ICPSR) marked a contrast to the norm. As a membership-based consortium, data orders were placed through and received by the university's Official Representative.

Print publications, such as *The Machine Readable Archives Catalogue of Holdings* (1981) or word of mouth were the only ways to discover the existence of data collections. An initial Canadian union catalogue, the self-described first annual edition of the Social Science Data Inventory, was produced in 1977 by the Data Clearing House for the Social Sciences. After the release of this catalogue, the Data Clearing House abruptly closed. A subsequent chill around developing national data infrastructure has been attributed in part to this closure.<sup>6</sup> This has been the source of much frustration for Canada's data professionals. Until the broader adoption of e-mail as an information dissemination tool, the creation of the Data Liberation Initiative (DLI), and the proliferation of web-based search tools, the landscape for finding data remained relatively unchanged. Proposals for the creation of a national data archive for Canada came to naught. Had such an agency been founded, the landscape for data discovery, sharing, and access would likely have been different.

## Peer-Group Organizations

In the 1980s, data professionals were less likely than now to have campus colleagues who shared their experiences of supporting data. Therefore, they had to find peers from outside their institution to collaborate with or to rely upon for assistance. The International Association for Social Science Information Services and Technology (IASSIST, <http://www.iassistdata.org>) made up of data producers in government and academia, data librarians, administrators, and researchers, was founded to provide these support mechanisms.

The inaugural meeting which led to the development of the organization was held in Toronto in 1974 concurrent with the 1974 World Sociology Congress.<sup>7</sup> The aim was to use the meeting to "bring issues of common interests to a joint gathering of social science researchers and data service providers."<sup>8</sup> The eighteen members of the first IASSIST Steering Committee represented twelve different countries. Their initial objectives were laid out and their activities were planned in the structure of action groups, including data organization and management, data archive development, data documentation, classification, process-produced data, and data acquisition. The leaders of the IASSIST committees realized that

standards were required for citing, formatting, cataloguing and referencing data files. Perhaps the best-known product of these groups was the *Working Manual for Cataloging Machine-Readable Data Files*, prepared by Sue A. Dodd, the U.S. chair of the Classification Action Group.<sup>9</sup>

During the inaugural meetings, Canada was represented by members from the academic community, the Public Archives of Canada, and Statistics Canada. The third IASSIST newsletter included reports of action groups on cataloguing and citation and on standards for data documentation, and a report describing a *Guide to Providing Social Science Data Services* presented at the May 1977 IASSIST conference in Toronto.<sup>10</sup> IASSIST has continued to influence the international profession of data librarians, researchers and producers for four decades with Canadian data librarians making significant national and international contributions to the organization.

The Inter-university Consortium for Political and Social Research (ICPSR), founded in 1962 at the University of Michigan, has long been linked with IASSIST. For example, one of ICPSR's earliest employees became the first president of IASSIST in 1976. ICPSR was created to make datasets available for reuse. The concept of reusing data, rather than the files being proprietary, was a foundational shift that led to the need to have data professionals at campuses who could serve to provide access to these products, technical expertise, and guidance. As described on the ICPSR website:

The impulse [of ICPSR] to break with proprietary tradition was both strange and much welcomed in the social scientific community. In hindsight, we can think of this data sharing motive as a prerequisite of the “scientific ethic” of verification, replication, and validation. In the early 1960s, though, the concept of giving access to all interested scholars to one’s basic (micro) data was so foreign as to be considered “revolutionary” ...likened to a violation of basic economic precepts: data were the scientist’s capital, and “they weren’t about to share their capital.”<sup>11</sup>

ICPSR has been an influential organization in the development of data professionals and data librarians in Canada. Two ICPSR membership consortia, CAP-PUL and OCUL/CREPUQ, represent 34 Canadian universities.<sup>12</sup> The twelve-person ICPSR Council includes two Canadian representatives. Canadians have participated in its educational program through courses and webinars in such topics as data curation and management, the ICPSR Summer Program, Official Representative Meetings, and mailing lists. ICPSR has served as a model and exemplar over the years for data service and delivery, and as a participant in and advocate of new projects such as the development of the Data Documentation Initiative.

## Impetus for Change within Canada

While Canadian data centres were still few in number and benefiting from organizations such as IASSIST and ICPSR, the academic community faced an unexpected crisis in November, 1984. Canada's Progressive Conservative government announced plans to cancel the quinquennial (1986) census. Less than a month later, the census was reinstated, but under a cost-recovery model unlike anything before in Canada.<sup>13</sup> In 1988, Statistics Canada announced the pricing model for census data. Western had purchased approximately 35 summary files (roughly 65 tables), and the three public use microdata files (PUMFs) from the 1981 Census for approximately Can\$5,000. For the 1986 Census, fewer tables would be available, each file would contain one table only, all three PUMFs were not expected, and the cost would be approximately Can\$110,000. The same scenario was taking place across Canada, as researchers began to realize the impossibility of affording to work with the Census. Institutions that had previously purchased census data realized that they would be unable to afford holdings comparable to 1981; therefore, a solution was needed to support academic research in Canada.

Laine Ruus had recently established a data library in Canada's largest university, the University of Toronto (UofT). Ruus proposed to the directors of the Canadian Association of Research Libraries/Association des bibliothèques de recherche du Canada (CARL/ABRC) that a consortium of its members be formed to purchase a single copy of all public 1986 Census data. UofT would undertake to copy and disseminate the files to each member of the consortium. Twenty-five<sup>14</sup> CARL institutions agreed to participate, although few were currently providing data services.<sup>15</sup> Without warning, librarians who knew nothing about data would be called upon by their institutions to provide access to and support for data.

"What do I do with this magnetic tape, and with what's on it?" was a question that began to echo throughout CARL/ABRC consortium member institutions. The most frequently designated individual to take on the role of data support was the government documents librarian, well-versed in print materials but frequently unaware of and ambivalent toward machine-readable data and their support. In 1988, at the Washington IASSIST meeting, Canadian attendees developed the idea of forming the Canadian Association of Public Data Users (CAPDU), an association whose objectives, were "to secure and promote efficient access to and use of public data in computer-readable form."<sup>16</sup> This objective meant training the trainers: experienced librarians within data centres undertook to train data neophytes, whether one-on-one or through conferences and meetings. Laine Ruus and Wendy Watkins (Carleton University's data professional) were fundamental in developing these early training programs. Beginning with rather informal gatherings of colleagues from across Canada, they developed into formal meetings held in conjunction with the Association of Learned Societies in Canada or IASSIST.<sup>17</sup>

CAPDU meetings provided the first extensive Canada-specific data training on Canadian files. CAPDU was strictly a volunteer organization: web server space, mailing list management, conference coordination, preservation of records, and other resources were (and remain) at the mercy of those able to provide them. CAPDU's importance in providing training in the early years of the expansion of data services in Canadian libraries cannot be underestimated.

Having established a consortium to acquire data from the 1986 Census, and with the cost-recovery model in place at Statistics Canada, the inevitable next step was to establish additional consortia for surveys such as the General Social Surveys of Canada, the Aboriginal Peoples' Survey, and the Health and Activity Limitation Survey. A second and larger CARL/CREPUQ consortium formed for the 1991 Census, again ramping up training needs. Ultimately, the term "consortium" became a problematic word among the data community in Canada.<sup>18</sup> Whether a file would be included in a consortium, what the cost for the particular consortium might be, and the time and effort involved creating the consortium all contributed to making the availability of data unpredictable and unequal across the country's academic institutions. Something better was needed to provide more equal access to data at a price that might be both predicted and budgeted year over year.

## The Data Liberation Initiative

The scope of this paper is insufficient to describe the establishment of the Data Liberation Initiative.<sup>19</sup> After tremendous effort by Ernie Boyko of Statistics Canada and Wendy Watkins of Carleton University, and community buy-in on various levels (government, Statistics Canada, research councils and universities), a five-year pilot project was launched in January 1996. Over 50 Canadian universities joined this project, paying a subscription fee for access to Statistics Canada data including microdata files and aggregated data products.

Immediate and ongoing training was needed to make librarians aware of the range of data provided in the DLI collection, to justify the membership cost, and to provide basic instructions to members who had not participated in the CARL Consortia. Again, institutions were calling upon untrained librarians to act as the local official contact for the DLI and as the hub through which Statistics Canada data would be provided on campus. Individual institutions had to determine the level of service that they might offer. The most basic service was to download data files from Statistics Canada and to provide them directly to users, while at the high end an institution might maintain a statistical support centre and operate a local online data retrieval system. To help new staff learn the tricks of the trade, the DLI's External Advisory Committee (EAC) established an Education Committee to report to the EAC. The Education Committee initiated and has in the years since continued to conduct annual training in each of the four regions of Canada. It established the practice of conducting "boot camps" for new data providers,<sup>20</sup>

and developed a first version of what has developed into the *DLI Survival Guide*.<sup>21</sup> On a regular basis the regions have gathered together for national training sessions, with members learning from each other and making useful contacts with colleagues from across the country. These training sessions continue to showcase the *DLI Survival Guide* and highlight the type of data service that might be established at an institution.<sup>22</sup>

For the period from 1996 through 2007, training was generally offered to data librarians at regional and national DLI training sessions and at CAPDU meetings. An institution's DLI fees paid for one representative's travel to a DLI training session, and it became more difficult for data specialists to justify additional travel expenses for separate CAPDU meetings; consequently, CAPDU meetings came to be held in conjunction with DLI training.

## The DLI Training Repository

DLI training, delivered by data professionals from DLI-member institutions and by Statistics Canada subject matter experts, is conducted regionally rather than nationally in three out of four years, and topics vary. The Education Committee recognized the need to provide access to the content of all training sessions, and established a training repository (<https://cudo.carleton.ca/collection/dli>). This repository contains PowerPoint slides, PDFs, and exercises for approximately 500 sessions that have been delivered at DLI training in the period 1997 to 2015. The repository is accessible by anyone interested in learning more about Canadian data or tools and tricks of the trade, and it is well-used: "in 2014, there were 1,676 sessions of the repository which accounted for 7,867 page views."<sup>23</sup>

At many institutions a single professional still provides data support, often devoting only a fraction of his/her time to data. The level of service delivered is impacted by local technical support, skill sets of staff members, and administrative support at the institutional level.<sup>24</sup> For these solo service providers, attending IASSIST, ICPSR or Canadian data meetings provides an essential opportunity to network, to share strategies for providing access to data files, and to consider possibilities for service expansion. Through these various training opportunities, data librarians learn techniques and tricks which they can take back to their local institutions to share with their colleagues and broaden the data culture.

## Creation of Online Data Delivery Systems

From the late 1980s into the 2000s, a number of Canadian institutions with technical support teams and experience in data developed tools for delivering data files to users' desktops. In most cases, these systems were developed at in-

stitutions where the data service was integrated with computing support. Some systems (e.g., Queen's QWIFS and the Tri-University Groups system) were developed on top of commercial statistical software packages. Others, such as Calgary's LANDRU, UBC's ISLAND, and CREPUQ's Sherlock, were hybrid systems delivering data and statistics. UofT deployed Berkeley's Survey Documentation and Analysis software (SDA). Western's data delivery underwent a number of changes. An early 1980's Cyber computer-based tape delivery system served as the back end of a VAX Fortran-based Network Data Library System (1986). The metadata entered there was transformed into static web pages for the web-based Internet Data Library System (IDLS, 1995). IDLS Version 2 used metadata transferred from IDLS into a searchable database to drive a new web-based system (2001). Finally, elements of IDLS Version 2 and Sherlock were integrated to create the bilingual Equinox Data Delivery System (2009). New data librarians typically received initiation in these various data delivery tools at DLI (or CAPDU) training sessions. Institutions without extensive local technical support could use these tools to access data, since the DLI license agreement allowed them to partner with other institutions to deliver data. It has become the norm to collaborate with peers from other institutions or to piggyback on services developed by other institutions.

For example, the Computing in the Humanities and Social Sciences (CHASS) unit at the University of Toronto obtains data from providers and creates an interface to those data that are offered to other institutions on a subscription basis. Universities across Canada may subscribe to one or more services from CHASS, the most popular being CANSIM (Statistics Canada's time-series database) and the Census Analyzer (a common interface to Canadian geographically-based census aggregated data). Other than CHASS, the system which enjoyed the largest adoption among other institutions were the UWO-based Internet Data Library System and its successor Equinox, which delivered microdata, aggregated data, and geospatial layers to users' desktops. As the DDI standard has spread, the Ontario Council of University Libraries (OCUL) deployed a Nesstar-based DDI-compliant system, <odesi> (<http://odesi.ca>), launched in 2008 as a common platform to deliver data to all Ontario universities. OCUL also launched a geospatial complement to data delivery, the Scholars Geoportal, in 2012. Coincidental to the retirement of the Equinox Data Delivery System in academic year 2014/2015, <odesi> enjoyed a surge in subscriptions from non-Ontario universities. Even more recently, in 2015, Statistics Canada launched two services, a Nesstar service that disseminates microdata to DLI users, and a Beyond 20/20 Web Data Server delivers aggregated data and geospatial files.

## Broadening the Data Community

As more Canadian institutions subscribed to or hosted a user-friendly data delivery service, the number of librarians and library staff who were providing reference service for using data files grew.<sup>†</sup> Training opportunities for staff could be provided at local institutions as peer training or at regional or national DLI training events. Other specialized training opportunities became available, such as an OCU-sponsored two-day workshop on using DDI, and a three-day True North Science Boot Camp for Librarians focused on research data management in the sciences (<https://truenorth2015.ok.ubc.ca/>). In 2007, the University of Alberta Data Library launched the Winter Institute on Statistical Literacy for Librarians (WISLL), a workshop for librarians aimed at training them to a basic understanding of statistical concepts and support.<sup>25</sup> This training series has been an annual event since 2007, attended by librarians from across Canada, including public service librarians as well as data professionals.

## Access to Confidential Data

While the Data Liberation Initiative was providing access to microdata and to aggregated data, higher-level researchers continued to need access to non-anonymized data from Statistics Canada to do more robust analysis or to look at variables not included on the public files. To support this research, “in 1998, the Canadian Initiative on Social Statistics ... [recommended] the creation of research facilities to give academic researchers improved access to Statistics Canada’s microdata files.”<sup>26</sup> Support from the Canada Foundation for Innovation (CFI) and the Canadian Institutes of Health Research, as well as local institutional support, has resulted in the establishment of 26 Research Data Centres (RDC) or RDC branches at universities across Canada.<sup>27</sup> The RDC program provides access to non-anonymized files to authorized users within a secure environment. Access to the RDC network is governed by a rigorous project approval process, which includes a police background check of the applicants for access. All analysis of RDC datasets is conducted within a secure data enclave constructed and paid for by the hosting university. An onsite Statistics Canada employee reviews researchers’ analysis to ensure no breaches of confidentiality.

In 2012 Statistics Canada launched the subscription-based Real Time Remote Access (RTRA) program as a complement to and partial substitute for establishing a RDC.<sup>28</sup> It offers simpler and faster project approval process, but provides less

---

† Although there are only 79 DLI institutions, messages to the DLI mailing list today are sent to 323 recipients. As a further example of the broadening of data responsibilities, from discussions with Walter Giesbrecht, and looking at the ads, York University advertisements for reference librarians now generally include a preference for experience with data.

robust access to the data because hands-on access to the data is not provided. Instead, a user sends a SAS syntax file to Statistics Canada, where it is executed. The descriptive statistics created by the analysis are automatically checked for confidentiality and are returned to the user if no breach of confidentiality is found.

The data librarian faces challenges in supporting users who wish to use secure data through either system. RTRA requires the use of SAS, a statistical programming language that is not customarily used by librarians. With either RTRA or the RDC, librarians face barriers in knowing what is contained in the system; while the survey questionnaires are publicly available, no variable-level search tool exists to quickly or easily identify files of interest to researchers.

A concern that has been voiced by data librarians since the launch of the RDC program was a fear that the number of PUMFs produced for the DLI program would be reduced since they are relatively costly to create, and since “real” researchers could use the files in the RDCs. That concern has been magnified by the deployment of RTRA and by budget cuts undergone by Statistics Canada. To some extent this fear has been realized, as a number of files (e.g., Longitudinal Survey of Immigrants to Canada, Community Noise and Health Study, Maternity Experiences Survey) have never been provided to the DLI community as public use files.

Given that Statistics Canada surveys may be released in any, all, or none of these forms, data librarians must be aware of the existence and availability of data and statistics at all levels in order to direct users to the most appropriate resource. The table below is based on the one contained in the current DLI Survival Guide, with the addition of the Real Time Remote Access Service.<sup>29</sup> It shows the Statistics Canada services with which a data librarian should be acquainted.

**Table 20.1.** Data Librarian Continuum of Access. (Reprinted with permission from DLI Survival Guide)

	<b>Open Statistics</b> ← → <b>Restricted Data</b>					
Service	Statistics Canada Website	Depository Services Program (DSP)	Data Liberation Initiative	Custom Tabulations	Real Time Remote Access and Remote Job Submission	Research Data Centres

As the number of Statistics Canada’s services increase, the need for a common search tool for these services becomes more critical. Librarians need to find the most appropriate information for users and should not omit any service from consideration, whether through neglect, oversight, or lack of familiarity. A common search tool should be made available to end users and the public in reflection of the government’s move toward open data, to provide them with self-mediated access to information. They then can approach their libraries knowing what they wish to use and receive instruction in how to use it.

## Reflections on Opportunities and Challenges

Canada's academic sector has responded to many challenges around the delivery of data and of training. In this section, we speculate on issues that academic data librarians will likely face in the future. Canada's municipal, provincial and federal governments are increasingly embracing the open data movement and creating their own data catalogues and search engines.<sup>30</sup> While many more data files become available, finding them and determining their comparability will prove to be an issue that confronts us. Additionally, government open data sites are not always maintaining preservation copies of data files: the Government of Canada site indicates that you may "search open data that is relevant to Canadians,"<sup>31</sup> but what is relevant to Canadians may not be the same as what is relevant to Canadian researchers interested in tracking phenomena over time. Data librarians and researchers will need to try to influence government policy to ensure that data are not lost when overwritten by new, "more relevant" versions of the same measures.

Canadian academic funding agencies are moving along similar paths. On February 27, 2015, a new open access policy for research was announced: "all peer-reviewed journal publications funded by one of the three federal granting agencies to be freely available online within 12 months."<sup>32</sup> However, this policy does not require "NSERC or SSHRC grant recipients to make their data openly accessible or archived at this time."<sup>33</sup> It refers users to the SSHRC Research Data Archiving Policy, which encourages users to deposit data within two years of a project's completion,<sup>34</sup> but this recommendation has in the past been largely ignored by researchers, as there is no enforcement or follow-up.

On July 17, 2015, a long-awaited *Tri-Agency Statement on Principles on Digital Data Management* was released in draft form for feedback.<sup>35</sup> The draft declares that "data management planning is necessary at all stages of the research project lifecycle, from design and inception to completion."<sup>36</sup> It proposes responsibilities for researchers, research communities, research institutions, and research funders. Academic institutions and librarians are charged with responsibilities, including providing access to "repositories ... that securely preserve, curate, and provide continued access to research data," providing researchers with "guidance to properly manage their data," and to "promote the importance of data management."<sup>37</sup>

As in the past, institutions will rely on collaborative efforts or borrowing from the work of others. The CARL Portage project to develop a community of practice for research data management in Libraries is the leading example, which deployed the Data Management Plan Builder created by the University of Alberta nationally as DMP Assistant (<https://portagenetwork.ca/>). Providing RDM services entails activities not currently undertaken at many Canadian universities, such as ensuring file integrity and migration from software platforms or storage media. Best

practices will need to be established and shared should these support tasks devolve upon data librarians and other public service librarians.

Unless we wish to rely on Google as the default global search engine for Canadian data, an agency will need to develop a comprehensive search engine for data or for data depositories.<sup>†</sup> This will become even more important as RDM becomes the norm within Canadian universities and research is deposited for long-term access into more and varied repositories.

Canadian data librarians must remain vocal and informed defenders of the Public-Use Microdata Files (PUMF) program at Statistics Canada because PUMFs are the key tool for training users of the future. We need to advocate for the creation of a centre of excellence for the creation of PUMFs within Statistics Canada to improve their quality and documentation, speed their creation, and reduce the cost of creation. DLI contacts echoed this concern at the National DLI Training Day in 2014 by calling on the DLI to initiate a system of quality control for SPSS data set descriptions before distributing them to the user community.

In addition, Canada's academic data librarians were active in the campaign to reinstate the mandatory long form census, which was canceled for the 2011 cycle. On November 5, 2015, the community rejoiced, as the newly elected government reinstated that program in time for the 2016 cycle, and has promised autonomy from ministerial interference for Statistics Canada.

IASSIST was founded in 1974 to address data organization and management, data archive development, data documentation, classification, process-produced data, and data acquisition. As shown in this chapter, Canadian data librarians have been responsive to changes in the data environment in Canada, developing, using and promoting collaborative tools, and developing and delivering training programs to their colleagues. Consequently, today's data librarians have access to many more tools and support for their activities than in the pre-DLI era. However, the issues around which IASSIST was found continue to face the data librarian in Canada and abroad.<sup>‡</sup>

---

<sup>†</sup> An international database similar to OpenDOAR (the Directory of Open Access Repositories, <http://www.openoar.org/>) would be ideal.

<sup>‡</sup> Maps showing the expansion of Canadian data services are available on the companion web site, <http://databrarianship.wordpress.com/> The authors offer our thanks to Christine Homuth and Brent Larue, of the Western Map and Data Centre, for their advice in preparing maps showing the expansion of Canadian data services. Thanks are also due to the DLI staff, particularly Renée Rocan, who located a spreadsheet that allowed us to map the DLI program membership in 2002 and 2015. Finally, we thank the Canadian Association of Research Libraries' Manager of Administration and Programs, Katherine McColgan, for delving through CARL's records to give us a list of the institutions that participated in the two census consortia.

1. Laine Ruus, "The University of British Columbia Data Library: An Overview," *Library Trends* 30, no. 3 (Winter 1982): 397.
2. Barton Clark, "Social Science Data Archives and Libraries: A View to the Future," *Library Trends* 30, no. 3 (Winter 1982): 505.
3. Laine Ruus, "Training of Data Services Professionals: Past, Present and Future," *Library Trends* 30, no. 3 (Winter 1982): 456.
4. Ruus, "The University of British Columbia Data Library," 398.
5. Slavko Manojlovich, "A Library-based Reference Service for Machine-Readable Census Data: The Canadian Experience." *IASSIST Newsletter* 3, no. 3 (1979): 55–64.
6. Chuck Humphrey, "Preserving Research Data in Canada: The Long Tale of Data Blog," accessed April 24, 2015, <http://preservingresearchdatainacanada.net/tag/data-clearing-house-for-the-social-sciences/>
7. Margaret O'Neill Adams, "The Origins and Early Years of IASSIST." *IASSIST Quarterly* 30, no. 3 (2006): 5–13. <http://www.iassistdata.org/downloads/iqvol303adams.pdf>.
8. *Ibid.*, 7.
9. *Ibid.*, 11.
10. "Report of the Joint Canadian-United States Action Groups on Classification," *IASSIST Newsletter* 1, no. 3 (1977): 8–9; "Standards for Data Documentation," *IASSIST Newsletter*, 1, no. 3 (1977): 22; "Organization of Data Archives," *IASSIST Newsletter*, 1, no. 3 (1977): 23.
11. Erik W. Austin, *ICPSR: The Founding and Early Years*, accessed May 21, 2015 <http://www.icpsr.umich.edu/icpsrweb/content/membership/history/early-years.html>.
12. Queen's and UBC were listed in ICPSR's *Annual Report* for 1964–65; Western in the 1965–66 issue. Thanks to Mary Vardigan at ICPSR for tracking down this information for us.
13. Bill Curry, "Census Debate is Nothing New," *The Globe and Mail*, August 16, 2010, accessed April 26, 2015 <http://www.theglobeandmail.com/news/politics/census-debate-is-nothing-new/article1377285/>.
14. The figure of twenty-five is drawn from Laine Ruus, "Semi-permeable boundaries among institutions: The Canadian Scene(s)," (presentation at the *International Association for Social Science Information Services & Technology*, Tampere, Finland, 2009), [http://www.fsd.uta.fi/iassist2009/presentations/A2\\_Ruus.ppt](http://www.fsd.uta.fi/iassist2009/presentations/A2_Ruus.ppt).
15. Ernie Boyko and Wendy Watkins, "The Canadian Data Liberation Initiative: An Idea worth Considering?" *International Household Survey Network Working Paper No.006*, November 2011, accessed April 29, 2015 <http://www.ihsn.org/home/sites/default/files/resources/IHSN-WP006.pdf>; Chuck Humphrey, "Collaborative Training in Statistical and Data Library Services," *Resource Sharing & Information Networks*, 18, no. 1–2 (2005):170, doi: 10.1300/J121v18n01\_13.
16. Canadian Association of Public Data Users, *By-Laws*, accessed April 29, 2015 [http://equinox.uwo.ca/capdu/CAPDU\\_by-laws.html](http://equinox.uwo.ca/capdu/CAPDU_by-laws.html).
17. Canadian Association of Public Data Users, *Meetings*, accessed April 29, 2015 [http://equinox.uwo.ca/capdu/CAPDU\\_meetings.html](http://equinox.uwo.ca/capdu/CAPDU_meetings.html).
18. Ernie Boyko and Wendy Watkins, "Data Liberation in 2004: How Did We Get Here?" (presentation at the *Data Liberation Initiative* Regional Training, Ontario, Canada, 2004), <http://hdl.handle.net/1873/232>.
19. For a thorough history of the development of the DLI, see Boyko and Watkins, "The Canadian Data Liberation Initiative. An Idea worth Considering?"
20. Boyko and Watkins, "The Canadian Data Liberation Initiative. An Idea worth Considering"?, 6.
21. Data Liberation Initiative Education Committee, *DLI Survival Guide*, last modified December 1, 2015 <http://www.statcan.gc.ca/eng/dli/guide/toc/3000279>.
22. Vince Gray, "Levels of Service," (presentation at the *Data Liberation Initiative* Regional Training, Ontario, Canada, 2004), <https://cudo.carleton.ca/collection/dli/search/levels%20of%20service>; Laine Ruus, "User Services in a Data Library," *IASSIST Newsletter* 4, no. 2 (1980), 29–33.

23. E-mail from Statistics Canada (Chantal Ripp) to DLILIST@statcan.gc.ca, May 13, 2015.
24. Vince Gray, “Levels of Data Service: The 2012 Ontario Context,” (presentation at the *Data Liberation Initiative* Regional Training, Ontario, Canada, 2004), <https://cudo.carleton.ca/dli-training/2766>.
25. University of Alberta, “Winter Institute on Statistical Literacy for Librarians (WISLL),” accessed May 18, 2015 <https://sites.google.com/site/wisll2014/home>.
26. Statistics Canada, “The Research Data Centres (RDC) Program,” accessed May 18, 2015 <http://www.statcan.gc.ca/eng/rdc/index>.
27. Canadian Research Data Centre Network, “Where to Access Data,” accessed May 18, 2015 <http://www.rdc-cdr.ca/find-a-rdc>.
28. Statistics Canada, “The Real Time Remote Access (RTRA) system,” accessed May 18, 2015 <http://www.statcan.gc.ca/rtra-adtr/rtra-adtr-eng.htm>.
29. Data Liberation Initiative Education Committee, “DLI Survival Guide.”
30. Statistics Canada, “Open Government Across Canada,” accessed May 18, 2015 <http://open.canada.ca/en/maps/open-data-canada>.
31. Statistics Canada, “Open Data,” accessed May 18, 2015 <http://open.canada.ca/en/open-data>.
32. Government of Canada, “Minister Holder Announces New Open Access Policy for Research,” accessed May 20, 2015 <http://science.gc.ca/default.asp?lang=En&n=415B5097-1>.
33. Government of Canada, “Frequently Asked Questions,” *Open Access*, accessed May 20, 2015 <http://www.science.gc.ca/default.asp?lang=En&n=A30EBB24-1#wsCD7C9989>.
34. Social Sciences and Humanities Research Council, “Research Data Archiving Policy,” last modified April 8, 2014 [http://www.sshrc-crsh.gc.ca/about-au\\_sujet/policies-politiques/statements-enonces/edata-donnees\\_electroniques-eng.aspx](http://www.sshrc-crsh.gc.ca/about-au_sujet/policies-politiques/statements-enonces/edata-donnees_electroniques-eng.aspx).
35. Government of Canada, “Draft Tri-Agency Statement of Principles on Digital Data Management,” last modified July 20, 2015 <http://www.science.gc.ca/default.asp?lang=En&n=83F7624E-1>.
36. Ibid.
37. Ibid.