
Creating a National Peer-to-Peer Training Program for Data Librarians in Canada

Introduction

The creation of the Data Liberation Initiative (DLI)² in 1996 opened a new channel of access to quantitative and spatial data files, making unprecedented amounts of Statistics Canada's data available for scholarly research and teaching through an affordable annual fee. The responsibility for the collection and provision of service for the DLI fell primarily upon academic librarians, the majority of whom were neophytes to the exciting world of data within an academic setting. This paper examines the training strategy developed in response to an urgent, Canada-wide, need for a new program of data dissemination³.

Background

The DLI was established as a result of negotiations between Statistics Canada, the Social Science Federation of Canada's Data Liberation Working Group and academic libraries in 1996. Initially presented as a five-year pilot project, an extremely positive evaluation during the fourth year of operation led to the transformation to an ongoing partnership program in 2001.

Under the DLI licence, Statistics Canada provides participating institutions with access to all of its standard data products, which includes databases, public use microdata files and geography files.⁴ As part of this contract, universities agree to make the data available to members of their communities and to guarantee that the data are used only for non-commercial teaching and research purposes. An important element of the DLI licence is that each institution must designate a local staff member as the DLI Contact who serves as the intermediary between Statistics Canada and his or her local institution. The DLI contacts are responsible for providing local access to the collection of DLI data products and for ensuring that the licence between their institutions and Statistics Canada is fully observed. In most cases, the DLI contact is an information-professional in an academic library.

The need to provide training for DLI contacts was identified as an early priority in the DLI pilot period. Without an established baseline of competencies among DLI contacts across all subscribing universities, local services would vary radically across institutions. Instead

by Wendy Watkins¹, Elizabeth Hamilton, Ernie Boyko and Chuck Humphrey

of being facilitators to access, DLI contacts could have become a bottleneck in providing local access to DLI data. The success of this initiative was highly dependent upon DLI contacts developing a basic level of knowledge about the DLI collection and acquiring a set of skills to disseminate data products to local patrons.

The Training Challenge

While DLI data resources greatly increase the potential for empowerment and insight into Canadian society, they also create new challenges for the librarians and information professionals who are confronted with the task of organising and supporting the access to this material. Prior to DLI, there were fewer than a dozen data libraries in the country and these were minimally staffed. For the new staff assigned to administer DLI, the challenges included: quickly acquiring skills to manage this new licensed resource; understanding the collection and its use; and developing skills to aid patrons in working with these data resources.

The challenges were a bit staggering. In addition to the paucity of experienced data librarians across member DLI institutions, many of the DLI contacts already bore multiple responsibilities in their libraries and had little time for yet another professional development venture on the job. Some came to the task with self-confessed statistical literacy deficits, and were intimidated by data and its technology. Few had resources sufficient to launch a full-service data library immediately and were dependent on finding colleagues both in and outside of the library to help build a local service model. Indeed, one of the training challenges was to develop a curriculum flexible enough to provide for tremendous disparities in local environments. Furthermore, the pre-existing data library community was small, stretched across a continent, and had to operate with two official languages (English and French). These vast distances introduced further challenges. For one thing, training costs were a real barrier unless the program was built in such a way as to encourage those without expertise to attend training sessions.

While the importance of building regional strengths was seen as part of the long-term solution to offering data services across the country, the identification of a common curriculum of essential skills appropriate to the different environments of small, medium and large libraries became an initial priority. This curriculum had to recognize the different starting points for DLI contacts within the training program. An initial curricular strategy was designed with these realities in mind.

A basic level of data service skills (core competencies) was first articulated as part of the curricular strategy. This training would be considered the entry level for staff supporting DLI data and would apply to all participating institutions, regardless of institutional size. More advanced training would build upon this basic level. More recently, the curriculum has been revised to address wider issues of data literacy (see **Appendix A** for the latest version of the Curriculum.)

All training was conducted from a ‘service’ perspective, that is, from a point of view focusing on the clientele of DLI data. The purpose of this training was to prepare data services staff to assist their user community with DLI data. This was an integral factor in the design and implementation of the training program. (See **Appendix B** for an overview of the DLI Training Principles and Implementation process.)

Following the drafting of the curriculum and training principles, considerable effort was invested to ready the scene for national training of the type envisaged. The first task was to build the core team of trainers for the first wave of training. With only a few seasoned data librarians at the outset, additional trainers had to be recruited from the existing Canadian library community and regional contacts. The maxim for this team became “as one learns, one will teach.” This principle filtered down to the current training delivery model.

Because the initial wave of national training was to be conducted in four regions of the country, a comprehensive manual was developed in both official languages. This served as both a training and future reference resource. This manual was distributed in a massive three-ring binder and included everything from technical notes to in-class and take-home exercises. The trainers, new and experienced alike, grew into a supportive community as they gathered in the spring of 1997 to prepare for the delivery of four regional workshops.

This community-building approach worked. The new trainers became better acquainted with the veterans, setting up a safety net for those trying to learn and teach data skills at the same time. All were encouraged and, indeed, energized by the retreat-like experience in developing data competency. Following the “Training the Trainers”

session, the team was ready to take the curriculum into the four regions of Canada, from coast to coast.

Central locations within the four regions were selected for training and, in the case of new trainers, a seasoned trainer was on-site to infuse the process with confidence. The seasoned trainers did not take a lead role. To do so would have undermined the role of the new trainers. But the experienced data colleagues performed trouble-shooting with the new trainers and clarified issues on data file verification processes, data parts, and elusive bits and bytes.

Among the factors contributing to the success of DLI training, two equally important conditions applied: one was to develop trainers in the regions who became recognized by their peers as mentors; the second was to get participants to training venues. Within six months of the inception of the DLI Training Program, a core team of trainers delivered training in all four regions covering both official languages, and over 80 librarians experienced one of these three-day training sessions in basic data competencies.

How were such numbers achieved? Travel stipends were built into the budget to cover transportation costs and, equally importantly, library directors were contacted to ensure that individuals assigned with the task of administering the DLI Project were granted the time to attend. The focus on four regional workshops was critical to the long-term goal of the program of establishing a data community to support the emergence of a successful data culture in Canada.

The On-going Training Program

From the outset, it was recognized that a single data training experience would not suffice to build the type of data competencies required to sustain the DLI Project. The data collection itself is a dynamic one. Furthermore, the technology that supports data use is in constant flux. As with many learning endeavours, there is typically a period of confusion at the beginning. Confusion tends to be followed by clarity as basic concepts are grasped. From clarity, competence develops allowing a layering of detailed information on the fundamental concepts. Finally, confidence is achieved through practice and with the transmission of skills or information to other colleagues. The training program had to allow for this type of learning experience, as well as the knowledge that, as individual competencies were increasing, so too were local environments changing for these individuals.

Following the regional training sessions in 1997, the DLI External Advisory Committee (EAC) discussed mechanisms for continuing the training. The extension of the training has allowed the curriculum to include topics appropriate to the maturation of the service, such as in-class instruction to students, and working with their colleagues locally to assist patrons with using Statistics Canada

products across the information continuum.

One of the most striking outcomes of the regional experiences was the development of a sense of community among the participants as they struggled to understand new concepts and worked through their homework together over a meal. Notable too was the continuation of contact between participants and trainers after the workshop. This sharing of expertise is in part a fulfillment of the mentoring that was intended through the choice of regional trainers. Several discussions have taken place on the substitution of alternate teaching methods to the workshops but, while web-based training can increase the knowledge base, the face-to-face workshops keep the community vibrant, informed, and cohesive. In a profession where individual participants may feel locally isolated, establishing a broader sense of a community of data professionals is key to attracting and retaining high quality staff.

The EAC gave priority in the DLI budget to continue annual training workshops through administrative support and stipends for travel for all DLI contacts to attend the workshops. Another observation following the initial training “boot camp” was the importance of the involvement of the DLI Section in Statistics Canada in the training. Between training sessions, they become the problem-brokers for DLI contacts on the front line of library services.

DLI contacts have participated in a total of twenty-eight training programs (seven in each of the four regions) and in 2003 met nationally in conjunction with IASSIST in Ottawa. They have not only been exposed to virtually all aspects of data, content, software and service, but have also developed a sense of community through their regional, national and international meetings. This community revolves around the training sessions, mentoring developed at the training sessions, and other support mechanisms such as a newsletter (the DLI Update), a listserv, and the development of a strong central DLI Section within Statistics Canada

Evaluation and Future Developments

From the discussions on the DLI listserv and from the high turnout at each regional workshop, there is strong evidence that the delivery of data services has grown substantially and that we now have aware, committed, and skilful colleagues across the country positioned to assist users with data disseminated through DLI. Colleagues in the areas of the country with solid data services at the outset have perhaps moved more quickly to a point where training is now being offered by a new generation of data librarians, as well as the experienced data trainers. In addition, as data services become established, library directors are increasingly advertising for people with expertise rather than adding yet another hat to an already fully-employed professional.

What are the success factors of DLI Training? While the above discussion focussed on some of the key principles, other factors contributing to this success can be identified. Some of these were through design, while others were by circumstance.

Focus: A clearly defined strategy was held that involved a target audience, a common core curriculum, and a set of priorities relating to a regional focus, bilingual training, and a commitment to public service.

Involvement: The training came from within the community. Statistical software providers did not teach SAS, STATA or SPSS. Rather, this instruction was provided by colleagues in the field who are sensitive to the starting points of our colleagues and who can explain the role of the software in the broader picture of data services. Similarly, discussions on Statistics Canada products have not come from divisions remote from end users. Rather, collection issues have been dealt with by those working on the reference desk who can succinctly answer the questions as to the purpose and use of data files associated with a product. Interestingly, the definition of the community has an elastic nature and is expanding to include as colleagues individuals from Statistics Canada author divisions who understand the DLI training program’s goals.

Relevance: The window of opportunity was limited, forcing the DLI Training Committee to act expediently rather than with a focus on perfection. Launching such a large-scale training program so quickly after the need was identified meant that the direction of subsequent workshops has been toward keeping pace with change. It has given the workshops a vital focus, building upon basic skills with timely and critical knowledge for data librarians.

Equality and Accessibility: The training is based on equitable access. Each region is served on an equitable basis and the training is designed for maximum accessibility by the provision of travel support. Each participant is treated as a colleague, regardless of her starting point in the curriculum or in the size of her home institution. Furthermore, those trained are given the tools to teach others, both colleagues and patrons alike.

Renewal: The first wave of training was followed by a reassessment, at which point training became part of the DLI strategic plan with its own budget component. Each region is open to the participation of DLI contacts from outside the region and, as importantly, in providing opportunities for trainers to continue to learn from other trainers. Following several years of gradual turnover in positions, the second and third “boot camps” for new DLI contacts were offered in 2001 and 2004. Another “Train the Trainer” initiative was also offered in 2004 with a focus on leadership and renewal as well as skills.

Commitment: A commitment was made by DLI to the library directors of member DLI institutions to provide training for their staff. In return, the expectation was made that the directors would support their DLI contacts in participating in this training. Consequently, participants come from institutions with an understanding that training is integral to DLI. This commitment to training has been the acceptance of responsibility at the individual and institutional level for the success of the program.

After seven years of DLI workshops, new challenges are emerging. Libraries are entities in constant change and there has been turnover in the DLI contact community. New materials and new access channels are opening up and peer-to-peer training is moving down to mainstream reference services. We are making new discoveries about statistical literacy and data service delivery that will be reflected in changes in the content of the curriculum. And, regrettably, some of the trainers are now approaching stages where they are making life changes because of retirement or relocation.

There are also new opportunities. With the core level competencies identified and basic training well underway through the workshops, other learning methods and technological solutions can be added to the training program and used within the DLI community. While the regional approach to training was adopted to build strengths in all areas of the country and allowed the training to address common interests of the different communities, there is a need to expand the world of talent and expertise within these regional data communities. The 2003 IASSIST meeting was held in Ottawa and was the venue for Canada's largest DLI training event ever. All DLI contacts were provided with travel support to attend national and international data workshops. Participants were able to choose up to four sessions from a total of twelve, ranging in level of expertise from novice to cutting-edge.

Since IASSIST is held in Canada once every four years, the plan is to organize a national training event each time IASSIST meets in Canada. This will be done in conjunction with the annual, regional workshops and will ensure that DLI participants have an opportunity to meet those from outside their region and the country who share many of the same trials and triumphs that they face.

Conclusions

Data Liberation began as a pilot project and was not established as an on-going program until after it underwent an official independent program evaluation. The role and importance of training was clearly a key factor in the positive evaluation of the program. It is significant to note that this was underscored by each group of stakeholders approached by the evaluators -contacts, managers, academics and Statistics Canada personnel.

Continued investment in training has become a given for Canada's data community. As the community grows in expertise, more and more members are able to take part in the enterprise as trainers, thus ensuring that those leading the sessions are a renewable resource.

The success of DLI in Canada has opened a new chapter in library service and sets the stage for creating a more numerate society. While we have yet to address fully the concerns raised by Dr. Paul Bernard, we are well on the way.

Concerning such issues, the public must have appropriate knowledge and not only hypothetical access to the data. Paradoxically, indeed, contemporary societies offer a wealth of information, but workers and citizens can be totally mystified, surrounded as they are by data whose flow and codes they do not master⁵

The challenge that we faced evolved as technology and information practices changed. The solutions were found within the community itself. The initial talent pool was carefully extended to trainers and then on to the broader community. Through the use of peer-to-peer instruction, a core set of competencies has been established throughout data library services in Canada. This in turn has produced new strengths as this new generation of data service providers has enriched and been enriched by the wider community through IASSIST. Clearly the next challenge is to expand the peer-to-peer approach to the larger community of data users.

Appendix A

DLI Curriculum

	Knowledge	Skills	Attitudes
Statistics and Data Literacy	<ol style="list-style-type: none"> 1. Understanding the framework of statistics and data 2. Understanding the Continuum of access 3. Understanding the uniqueness of data as a medium 4. Understanding the methods of data collection 	<ol style="list-style-type: none"> 1. Recognizing a data question 2. Interpreting data documentation 	<ol style="list-style-type: none"> 1. Overcoming anxiety 2. Nurturing sharing and open access 3. Nurturing preservation
Content	<ol style="list-style-type: none"> 1. Knowing where and how are statistics gathered (STC) 2. Knowing the collection (DLI) 3. Knowing about other data collections 	<ol style="list-style-type: none"> 1. Finding 2. Accessing 3. Using 4. Sharing 5. Re-purposing (creating new data from old) 	<ol style="list-style-type: none"> 1. Advocating access to, openness to and use of statistics and data 2. Being tenacious 3. Being creative and bold

Tools	<ol style="list-style-type: none"> 1. Understanding the variety of options: data, software, output file types, post processing 2. Selecting appropriate tools 3. Developing knowledge of access tools 	<ol style="list-style-type: none"> 1. Using statistical packages 2. Using the web tools (access tools) 3. Understanding search tools 	<ol style="list-style-type: none"> 1. Being open to lifelong learning (re-learning tools) 2. Being curious and having the courage to attempt the new 3. Being positive towards change
Services	<ol style="list-style-type: none"> 1. Recognizing the variety of options for service models 2. Adapting a service model responsive to internal and external changes 3. Understanding the user/audience 4. Being aware of funding sources 5. Administering the service and DLI license 	<ol style="list-style-type: none"> 1. Conducting an environmental scan 2. Writing grant applications 3. Interpreting the DLI license 	<ol style="list-style-type: none"> 1. Holding a positive attitude toward service to a large community 2. Being a champion for data; one who is proactive about promoting data service 3. Being positive about consulting (data reference is a consultation) 4. Developing a data culture

APPENDIX B

DLI Training Principles

The following principles serve as guidelines for the overall training program sponsored by the Data Liberation Initiative project. These were developed by the National Training Committee in December 1996 and subsequently modified by the DLI EAC Education Committee in October 2003.

1. Training under this program is being conducted specifically for (1) DLI contacts at participating universities, (2) the staff who will provide services for DLI data at these institutions, and (3) Statistics Canada staff directly involved in the support of DLI.
2. Training will be provided to all of those eligible under the first principle through a variety of formats, including subsidized workshops that are delivered regionally.

3. The first training priority is to establish a basic level of data service skills for new DLI contacts. This training shall be considered the entry level required to DLI data. More advanced training will build upon previous levels. Priorities for advanced levels will be determined by the needs of those supporting data services and by the evolution of DLI.
4. Training priorities for Phase II will address varying levels of expertise and service within the DLI community. Special attention will be given to strengthening expertise in regions undergoing changes due to retirement or turnover in experienced DLI contacts. Special attention will be given to regions without a prior tradition or culture of data use or without a previous foundation in data services. This will entail strategies that identify the training and support of key individuals who will become recognized experts in their region.
5. All training will be conducted from a 'service' perspective, that is, from a point of view that focuses on the clientele of DLI data. The purpose of this training is to prepare data services staff to assist university clients with DLI data.
6. A global curriculum plan will guide the course content that is offered through this program. The DLI External Advisory Committee will be responsible for maintaining this plan and for periodically reviewing its content and direction.
7. Training will address concerns appropriate both to small and large institutions.
8. Training will be regionally based with regular national and international exposure when the opportunities arise.
9. Whenever possible, trainers will be recruited from the existing Canadian data library community with the expectation that those who are trained may some day be called upon to train others. This perspective operates on the principle that as one learns, one will teach.
10. Outreach will be organized for Library directors, the user community, STC survey managers, and other general public to communicate the importance of statistical and data literacy.

Notes

¹ Contact: Wendy Watkins, Data Centre, Carleton University, Ottawa, Canada K1S 5B6. Phone: +1 613 520-2600x8376. Internet: www.carleton.ca/~ssdata. Email: wwatkins@ccs.carleton.ca

² Watkins, Wendy and Ernie Boyko (1996) "Data Liberation and Academic Freedom" *Government Information in Canada/Information gouvernementale au Canada* 3, no. 2 (1996). At <http://www.usask.ca/library/gic/v3n2/watkins2/watkins2.html>

³ Adapted from a paper by Ernie Boyko, Elizabeth Hamilton, Chuck Humphrey and Wendy Watkins, presented to IFLA's 69th Conference, Berlin, August, 2003. This paper was also presented at the IASSIST Conference held in Madison Wisconsin in May 2004 in the session on "Developing Statistical Literacy: Think Globally, Work Locally.":

⁴ Examples of files included in the DLI collection are Census cartographic files at all levels of geography, Postal-code Conversion files, Canadian Community Health Survey, Volunteer Survey, Survey of Labour and Income Dynamics and the collection of Canadian General Social Surveys. A complete list may be found at http://www.statcan.ca/cgi-bin/spider/dli_list.cgi

⁵ Bernard, Paul (1992) "Data and Knowledge: Statistics Canada and the Research Community," *Society/Société*, May 1992, p. 22.