

## **INTRODUCING DATA INTO CANADIAN ACADEMIC LIBRARIES: THE STRAW THAT DIDN'T BREAK THE CAMEL'S BACK**

Wendy Watkins

Carleton University, Canada

Charles Humphrey

University of Alberta, Canada

wwatkins@ccs.carleton.ca

*A decade ago, Canadian universities and Statistics Canada embarked on a joint adventure called the Data Liberation Initiative. While this initiative greatly increased the availability of data for academic research and instruction, it also created new challenges for the librarians and information professionals who were confronted with the task of organizing and supporting access to these resources. This paper will examine the challenges involved in the introduction of data into the traditional academic library. The experience of Canada's Data Liberation Initiative (DLI) will be used to illustrate the types of issues involved and the approaches used in the Canadian context to address them. The paper will conclude by identifying the best practices that have emerged after a decade of experience in developing, refining and promoting the use of data as a means of strengthening the underpinnings for teaching and research.*

### **INTRODUCTION**

The creation of the Data Liberation Initiative (DLI) in 1996 opened a new channel of access to quantitative and spatial data files, making an unprecedented volume of data from Statistics Canada available for scholarly research and teaching through an affordable annual fee. The responsibility for supporting local access to the collection and providing service to DLI resources fell primarily upon academic librarians, the majority of whom were neophytes to the world of statistics and data. This paper describes the national training programme developed to assist those new to data services.

### **BACKGROUND**

The DLI was established as a result of negotiations between Statistics Canada, the Data Liberation Working Group of the Social Science Federation of Canada and academic libraries represented by the Canadian Association of Research Libraries. Initially accepted as a five-year pilot project, a strongly supportive evaluation during the fourth year of operation (Goss Gilroy, 1999) resulted in a permanently established programme 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. As part of this contract, universities agree to make the data available to members of their institution and to guarantee that the data are used only for academic teaching, research and planning purposes. An important element of the licence is that each institution must designate a local staff member as the DLI Contact who serves as the intermediary between the university and Statistics Canada. These 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 the majority of 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 during the DLI pilot period. Without an established baseline of competencies among DLI contacts, the level of local services would vary radically across all subscribing institutions and contacts would run the risk of becoming bottlenecks to access instead of facilitators to the local use of these valuable data resources. 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 research and instruction about Canadian society, they also create new challenges for the librarians and information professionals who are confronted with the task of organising and supporting access to this material. Prior to DLI, there were fewer than ten data libraries across the country and many of these were minimally staffed. For those newly assigned to administer DLI on their campuses, the challenges were at least threefold: (1) quickly acquiring skills to manage this licenced collection; (2) understanding the collection and its use; and (3) developing skills to aid patrons in working with these data resources.

At the beginning, many DLI contacts found these challenges intimidating. In addition to the paucity of experienced data librarians across member DLI institutions, many of the contacts already bore multiple responsibilities in their libraries and had little time for yet another on-the-job professional development assignment. Some came to the task with self-confessed statistical literacy deficits and were frightened of data and the corresponding technology supporting data use. 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. 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). Vast distances between institutions and corresponding travel costs to bring people together in one location posed a serious impediment to holding training sessions. A successful training programme had to overcome these barriers and to motivate those without data expertise to attend training workshops.

While the importance of building on regional strengths was seen as part of the long-term solution to offering data services across Canada, the identification of a common curriculum of essential skills appropriate to the different environments of small, medium and large institutions became an immediate priority. This curriculum had to be structured to accommodate the different starting points for DLI contacts. An initial curriculum was designed with these realities in mind.

A basic level of data service skills (core competencies) was articulated as the first part of the curriculum design. 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.

All training was conducted from a *public service* perspective; that is, from the point of view of supporting the users of DLI data. The purpose of this training was to prepare professionals in data services to assist their user community with DLI data. This training was not intended to create new statisticians or social scientists. Rather, the goal was to equip information specialists with the knowledge and skills to provide first-rate data services. This was an integral factor in the design and implementation of the training programme (DLI, 2005a).

After drafting the curriculum and training principles, considerable effort was invested in preparing the launch of a national training programme. A core team of trainers was identified to work within four regions of Canada. With only a few seasoned data librarians at the outset, additional trainers had to be recruited from regional contacts in the library community. The maxim of this team became "as one learns, one will teach." This principle filtered down to the current training delivery model. Peer to peer instruction has been fundamental to the implementation of the DLI curriculum.

Because the initial wave of national training was to be conducted in four regions, a comprehensive manual was developed in both official languages. This manual, which served as both a training and reference resource, was distributed in a large 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 proved to be very effective. 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.

Central locations within the four regions were chosen as training sites and, in the case of new trainers, a seasoned trainer was available for encouragement and to instil confidence. The seasoned trainers did not always take a lead role, which would have undermined the stature of the new trainers. But the experienced data colleagues offered trouble-shooting assistance to the new trainers and, when called upon, clarified technical issues.

The success of the whole training enterprise was not only dependent on the development of trainers in the four regions who became recognized by their peers as mentors but also in getting participants to training venues. Within six months of the inception of the DLI Training Programme, a core team of trainers had delivered training in all four regions and in both official languages. More than 80 librarians experienced one of these three-day training sessions in basic data competencies. A significant contributor to this high level of participation was the travel subsidies provided through the DLI External Advisory Committee (EAC) budget to cover transportation costs. Because the distances that people had to travel varied greatly, covering travel costs to and from the venues allowed everyone to participate. Equally important, library directors were contacted to ensure that individuals assigned with the task of administering DLI locally were granted the time to attend. The focus on four regional workshops was critical to the long-term goal of establishing a national community to support the emergence of a data culture in Canada.

#### THE ON-GOING TRAINING PROGRAMME

From the outset, it was recognized that a single data training experience would not be sufficient to build the type of data competencies required to sustain the DLI programme. Several factors came into play over the years:

- some DLI contacts have changed jobs consequently creating the need to bring new contacts into the fold.
- new institutions have joined DLI since the initial training programme. Their contacts obviously were in need of an orientation to data services.
- the DLI data collection itself is dynamic with new content being added continually. To stay current, contacts have needed at least refresher courses on new materials.
- the technology that supports the use of data has been in constant flux. For example, access to several statistical databases evolved from standalone PC software to Web services between 1996 and now. The introduction of these technologies has always created new training opportunities.
- as DLI contacts gained more experience on the job, they became aware of gaps in their data knowledge and skills, leading them to seek further training.

All of these factors pointed to the need for ongoing training and, following the initial regional training sessions in 1997, the EAC approved a continuing education programme.

An Education Committee, composed of regional training co-ordinators and EAC members, operated on an ad hoc basis until 2003 (DLI, 2005b). At that time, the EAC formalised the structure of its literacy programme by giving the Committee specific terms of reference and defining its membership, which is made up of the two training coordinators from each of the four regions and representation from the DLI Unit in Statistics Canada. The Committee is responsible for making recommendations to the EAC regarding the overall training mission of DLI and changes to its curriculum. This formalised structure represents a final step in institutionalising a data literacy programme within the DLI.

A total of thirty-six workshops (nine in each of the four regions) have been hosted since 1997 and an additional national training event was held in 2003 in conjunction with the International Association for Social Science Information Service and Technology (IASSIST) in Ottawa. Throughout this training, DLI contacts have been exposed to a comprehensive mixture of data, content, software and service. They have also developed a sense of community through these regional, national and international meetings. Notable too, has been the contact among participants and trainers after the workshops. 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 among the training coordinators about substituting workshops with alternative distance learning methods, such as web-based instruction, but 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 community among data professionals is instrumental to attracting and retaining high quality staff.

In 2004, the original group of trainers recognized the need to renew and strengthen the training programme by adding new blood in their ranks. These trainers identified individuals from their regions who they saw as potential instructors. A three-day Train-the-Trainers workshop was held consisting of leadership instruction and indoctrination in DLI training principles. Since then, all of the new trainers have taught in regional workshops, having nearly doubled the size of the training corps.

## SUCCESS FACTORS AND FUTURE CHALLENGES

Strong evidence exists that the delivery of data to researchers, students and instructors in post-secondary institutions in Canada has grown significantly as a result of DLI and that the DLI training programme has well prepared those who provide these data services. This evidence can be found in an external independent programme review conducted in 1999, in a number of discussions on the DLI email distribution list, in the positive evaluations received following workshops and in the high level of participation at workshops, which typically have an 80 percent or higher attendance rate. In addition, library directors are increasingly recognizing the importance of data services as part of the overall complement of library services. This can be observed in the advertisements of some new positions that include data expertise as a desired skill.

A variety of indicators has been identified by the EAC to monitor the performance of DLI, including some of the ones mentioned in the previous paragraph. After nine years of workshops, a few factors have been particularly prominent in explaining the popularity of the training programme. First, the target audience has been clearly identified from the beginning. This has not been a programme in search of participants. We knew who the DLI contacts were and worked closely with them in developing training sessions. Furthermore, training was organised regionally to deliver workshops a propos to institutions with similar contexts and to facilitate the building of regional support communities.

Second, a common core curriculum was developed that has been applicable to the daily tasks of providing a public service. The DLI Training Committee prepared materials focusing on the immediate application of data knowledge and skills without belabouring theory. This kept the material relevant to the work that DLI contacts have been facing and has given the workshops a crucial focus, building upon basic skills with timely and critical knowledge for data services. Because of their heavy work schedules, DLI contacts need to see the immediate benefits of training to justify the time spent away from the office attending DLI workshops. Without this tangible payoff from DLI training, we would have much lower attendance rates at workshops.

Third, training was conducted by practitioners through peer-to-peer instruction. Training came from within the community. Each participant has been treated as a colleague, regardless of starting point in the curriculum or in the size of home institution. This open acceptance of DLI contacts helped build the sense that they belong to a homogeneous community. Those trained were given the tools and encouraged to teach others, both colleagues and patrons alike. For example, statistical software providers did not teach *SAS*, *STATA* or *SPSS*. Rather, this instruction was provided by colleagues in the field who can explain the role of the software in the context of data services. Similarly, discussions on many data products have not come from Statistics Canada staff, who are often isolated from end users. Rather, these products have been presented by colleagues working with them on the data reference desk. Nevertheless, this community of data services professionals has embraced and involved individuals from Statistics Canada author divisions who understand and share the DLI training programme's goals.

Finally, leadership has been essential in the success of the DLI training programme. Guidance and support have come through the EAC, which has made training a funding priority in its annual budgets. As mentioned earlier, travel subsidies have enabled people to participate who otherwise would have been excluded because of travel costs. The Education Committee

consisting of regional training coordinators has provided vital direction and vision for the programme. These coordinators have carried the operational responsibility for the programme and have been absolutely instrumental in its success. Their core commitment to this training has been the lifeblood of the programme. For the long-term health of the programme, a succession plan has been started to renew the training coordinators group from time to time. Related to leadership is a sense of ownership and DLI contacts feel that DLI training belongs to them. To the extent that contacts hold this feeling, training will remain a high priority within the community.

Certain challenges will need to be addressed in the second decade of DLI workshops. Today's academic libraries are pressure cookers. New materials and digital services are being acquired at ever increasing volumes with expectations that current staff do more with less. This places greater demands on those supporting services in our libraries. As a result, the time to dedicate to data services is in strong competition with these new demands on professionals already carrying heavy workloads. This challenges us to ensure that the value of DLI training always remains high in the eyes of DLI contacts. We must continually communicate with DLI contacts the value in upgrading their knowledge and skills in data.

Another challenge in this next decade will be coordinating related training activities that share a common global goal, such as building a data culture in Canada, but are not in step. For example, the Social Sciences and Humanities Research Council supports four annual Data Training Schools (DTS) in Canada. These programmes have been funded to expand the quantitative analysis skills of Canadian researchers working with social data. The programme has been directly associated with the Research Data Centres in Canada because of the confidential data from longitudinal surveys that they hold. One of the objectives of the DTS is to teach advanced statistical techniques for analysing longitudinal data. Because the public use microdata in DLI are almost all from cross-sectional surveys and because DTS training has been directed toward data analysts, DLI and DTS training have operated in parallel. The unfortunate consequence of this outcome is that the organisers of these programmes have never discussed how their curricula might complement one another or how common resources might be shared. Rather than rely on serendipity as the force shaping the outcome of national data culture, coordinated planning among these training programmes is more likely to produce desired results.

## CONCLUSIONS

Goss Gilroy's independent evaluation of DLI in 1999 attributed the healthy growth of the programme in its early years to ongoing training. 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 in Canada's data community has widespread support. Such investment, however, has spawned even greater returns than initially expected. As the community of DLI contacts grows in expertise, these contacts become increasingly valuable partners in instruction and research at their member institutions. More than one DLI contact has become an important contributor to projects seeking grant assistance because their expertise in data is relevant to new research. Similarly, DLI contacts are being increasingly asked to provide instruction about data in credit courses because of their data knowledge. DLI training is now paying benefits beyond high quality data services through the expanded talents of its DLI contacts.

Even though a firm foundation has been established upon which future DLI training can take place, an appropriate question is whether other organisations should take more responsibility in preparing the foundational skills and in building basic competencies in data services. Without the involvement of library schools in data services training or large professional associations offering certification in this area, the DLI programme has been an essential bridge to cover this gap in Canada (Humphrey, 2005). Until there is movement by professional associations and library schools to change this situation, the collaborative training programme offered through DLI will continue to prepare professionals with the knowledge and skills that they require to support data services.

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 the concerns raised by Bernard (1992) completely, 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 master.

The challenges 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 generation of data service providers makes contributions on their local campuses and in the wider community through involvement in organisations such as IASSIST and IASE. To further this end, both the Data Liberation Initiative and the Administration Committee of IASSIST have embraced the promotion of statistical literacy as a major strategic direction. DLI now encompasses three meanings: Data Liberation, Data Library and Data Literacy Initiatives.

#### REFERENCES

- Bernard, P. (1992). Data and knowledge: Statistics Canada and the research community. *Society/Société*, 16, (p. 22).
- DLI (2005a). *DLI Training Principles: The Pedagogical Foundations of the DLI Training Program*. [www.statcan.ca/english/Dli/train-prin.pdf](http://www.statcan.ca/english/Dli/train-prin.pdf).
- DLI (2005b). *Data Liberation Initiative Education/Training Structure*. [www.statcan.ca/english/Dli/educ-struct.pdf](http://www.statcan.ca/english/Dli/educ-struct.pdf).
- Goss Gilroy Inc. (1999). *Evaluation of the Data Liberation Initiative*. [www.statcan.ca/english/Dli/eval.pdf](http://www.statcan.ca/english/Dli/eval.pdf)
- Humphrey, C. (2005). Collaborative training in statistical and data library services: lessons from the Canadian data liberation initiative. In W. Miller and R. Pellen (Eds.), *Libraries Beyond Their Institutions: Partnerships that Work*, (pp. 167-181). New York: Haworth.
- Watkins, W. and Boyko, E. (1996). Data liberation and academic freedom. *Government Information in Canada/Information gouvernementale au Canada*, 3.