Developing a data service in an undergraduate Library: Past, present and future

IASSIST Conference, Ann Arbor, Michigan, May 26, 2006

Suzette Giles
Data, Map and GIS Librarian
Ryerson University Library, Toronto, Ontario

May 26th 2006
Context

- Ryerson University is founded in 1948 as Ryerson Institute of Technology
- 1963 - becomes Ryerson Polytechnical Institute
- 1971 - Ryerson given permission to grant degrees (Bachelor of Technology and B. of Applied Arts)
- 1992 - New faculty required to do research
- 1993 - Ryerson Polytechnic University is established with full university status including a fully funded research role and power to develop graduate programs.
- 2000 - Ryerson launches its first graduate programs
- 2002 - Ryerson University becomes the official name
## Ryerson Library

<table>
<thead>
<tr>
<th></th>
<th>1995-1996</th>
<th>2003-2004(05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students</td>
<td>13,800</td>
<td>20,700</td>
</tr>
<tr>
<td>Library staff</td>
<td>65</td>
<td>75</td>
</tr>
<tr>
<td>Librarians</td>
<td>15 (aver. 30)</td>
<td>20 (aver. 31)</td>
</tr>
<tr>
<td>Print holdings per student</td>
<td>30 (aver.105)</td>
<td>32.8 (aver. 109)</td>
</tr>
<tr>
<td>% Univ budget</td>
<td>3.59%</td>
<td>4.16%</td>
</tr>
<tr>
<td>Grad programs</td>
<td>None</td>
<td>15+ (700 stud.)</td>
</tr>
</tbody>
</table>
Past: 1995

- The Chief Librarian drops hints that the library should have a “data librarian”. Librarians confused – what is a “data librarian”?
- I volunteer even though I am not sure what the term means.
- Upside – I get to go to Quebec City to the 1995 IASSIST conference. Understand one word in ten if lucky, but the people are great!
- Seconded to University of Toronto Data Library Services for 1996 as the librarian, Laine Ruus is going on sabbatical
- My Chief Librarian and Laine turn out to be old friends – have I been set-up??
Past: 1996

- I have about six weeks of intensive training with Laine. Making progress – now understand one word in five and maybe a few numbers!
- Fortunately Laine stays in Canada for most of her sabbatical
- A challenging year includes the mainframe computer being closed down with 48 hours warning!
- But also attend the ICPSR data librarian course
- Considerably wiser, I return to Ryerson in January 1997 to develop a data library service
1997: Starting a data service

- I find there is not much of a demand for social science data and little awareness of what is available
- Many faculty are doing qualitative research or have acquired the datasets they need from their own sources
- Most newer, research based faculty are in Engineering
- There are no graduate programs therefore no graduate students requiring data
- Data are just an added responsibility with no resources attached (space, staff, equipment)
- The service develops very slowly!
Past: 1997-2002  Library based

However some important things are happening:

• The Data Liberation Initiative has become a reality in 1996 and I am the representative for Ryerson.
• DLI provides micro and aggregate data and, very importantly, training, networking and a listserv
• 1996 Canadian census data are becoming available along with boundary files – GIS is in the air.
• Ryerson joins the Ontario/Quebec consortium for ICPSR
• I take on responsibility for Statistics Canada publications in the library. This is a good fit with DLI responsibilities
Past: 1997-2002

A number of things are done to raise the awareness of the resources we have:

• A “Research Guide” on Statistical data is developed for the library website

• A listserv is started for faculty that notifies them of new data products and workshops and seminars that might be of interest

• Workshops and poster sessions are presented at annual the Ryerson Faculty conferences on PUMFs, census data, etc.

• Data Services is written up in library publications

• Shamelessly link to University of Toronto Data Library Services website
Past: 1997-2002

• As there is little demand for microdata I concentrate on aggregate products from Statistics Canada
  – Make sure Statistics Canada publications are in the library catalogue
  – Install specialized CD-ROMs on computers in the Reference area and develop procedure manuals for searching them
  – Have a workstation with software, data and procedures for mapping 1996 census data
• Subscribe to QWIFS, a data subsetting and analysis service from Queen’s University, Ont.
• Subscribe to University of Toronto’s SDA service when it becomes available
Past: 1997-2002

I also:

• Apply for any funding I can find – manage to upgrade my computer and printer

• Hire students under an Ontario work study program to install the CDs and write the procedures. The library only pays 25% of the cost

• Acquire geospatial data as it becomes increasingly available and relevant to programs

• Lobby – bringing up support for data every time one time funding or long term planning is mentioned. I make sure the service gets a mention in planning and space documents. Librarians start rolling their eyes.
Present: 2003-06
Reaching out to the university

- In 2000 a new Chief Librarian is appointed who lobbies tirelessly for increased resources for both staff and material and for improved space.
- By 2003 things are changing as funding for collections increases each year.
- And there are also new positions - librarians increase from 16 in 2000 to 22 in 2005.
- 6 new support staff positions including a fulltime data and GIS technician are announced in late 2003.
Present: 2003 - 2006

By 2003 the data service needs a higher profile:

- A website needs to be developed to support both social science and geospatial data
- Geospatial data is becoming more and more available through various agreements but the collection is in my head and a set of file folders.
- An index to our print map collection should also be made web accessible rather than printing out 1000 pages every summer
- Because some new positions are not yet filled I get funding for summer students
Present - 2003

• After surveying other Ontario university library sites we decide to do what University of Toronto is doing – deliver geospatial files over the web where allowable.
• Server space, access to web server software and expertise are required.
• The Library has neither a server, the software nor the expertise to support this project.
• Unexpectedly Computer and Communication Services are willing to give us 500 gigabytes of space on a new server and to allow students to download data from it.
Present - 2003

- CCS also helps the students to learn the software (Dreamweaver and Coldfusion) and to troubleshoot problems
- Over the summer the Research Guide is extended and updated and forms the Social Science data part of the site
- The Map Index is transferred to a Microsoft Access database and can be searched over the web
- A third section of the site is for Geospatial data
Present – 2003
Geospatial data

- An Access database forms the core of the geospatial data search capability
- Geospatial data, where the licence agreement permits, is made available by direct download from a secure site.
- Otherwise data can be requested by filling in an on-line form
- Data can be delivered via CD-ROM, memory stick or a secure, password protected website, once a licence agreement has been signed
- By October 2003 there is the beginnings of a Map and Data Resources (MADAR) website linked from the Library’s home page.
Present: 2003 - 2006

- A university fund raising campaign begins to make progress and all areas are asked to submit funding proposals that will attract donors.
- The library’s submission for funds includes a data and GIS area, which is put low on the priority list by both the library and the university.
- However the library is fortunate to get a donor who gives $1 million to refurbish the main floor and to develop an Information and Learning Commons.
- The Information and Learning Commons will include a Geospatial, Map and Data centre!
Present – University based

• Lobbying has paid off. I have more than I ever expected – staff, space and equipment
• Work on the ILC begins in the summer of 2004
• The Geospatial, Map and Data centre houses the map collection, 4 well equipped computers and an office for the GIS and Data technician
• The technician begins January 2005
• It is possible to make more files available over the web and to provided support to faculty and students in selecting and using geospatial resources
Old Reference area
Old Reference area
Information and Learning Commons
Information and Learning Commons
View of old map collection
Geospatial, Map and Data Centre
Geospatial, Map and Data Centre
Present - 2006

• In 2005-2006 there are over 500 specialized, in person, reference questions at the Geospatial, Map and Data Centre
• Over 16,000 users accessed and/or downloaded our secure files
• With the web presence, a dedicated space and full time staff we have become a true service point that reaches beyond the library.
Future – Beyond the university Possibilities and challenges

• Numeric/ Statistical data and geospatial data are facing similar challenges in the Ontario university environment
• Once the cost of obtaining data was a major hurdle, now the quantity can be a problem
• Each institution acquires, stores, manipulates and makes the data available to users
• There is duplication of effort
• The support and level of service varies considerably from institution to institution depending on resources and priorities
OCUL – Province wide

• The university libraries in Ontario form a group called OCUL (Ontario Council of University Libraries) that discusses issues of joint interest

• A major project is Scholars Portal which is intended to deliver a set of services to all Ontario Universities

• It began with acquiring group access to periodical databases and delivering the content to the university libraries
Overview of Scholars Portal*

– a set of best-of-breed applications tightly-coupled through standards-based linking
– supports **search**, **browse**, **save**, and **request** services in a simple, intuitive way
– services are presented and delivered through each library’s web portal
– infrastructure and content are managed centrally

* From a Scholars Portal presentation on the OCUL website
Key Message

• For the researcher Scholars Portal provides a single point of entry into an integrated and inter-connecting environment of high quality scholarly resources and sophisticated user focused services that enables:
  – rapid access to resources, services and tools
  – long term stewardship, archiving and persistence

* From a Scholars Portal presentation on the OCUL website
Scholars Portal Architecture

SEARCH Illumina

REQUEST RACER

SAVE RefWorks

BROWSE Science Server

Get It! OPENURL

Other A&I

Other Full-Text
Scholars Portal for Data?

Over the last few years the idea has emerged that perhaps numeric and geospatial data could be delivered in a similar fashion.

Two groups that meet under the auspices of the OCUL Directors are the map librarians (OCUL Map Group) and data librarians (DINO - Data in Ontario).

Both groups have been discussing the concept of centralized storage/archiving and delivery of data.

Several developments have made this approach seem possible for social science data.
A number of universities have developed their own data extraction and analysis tools but sustainability is a concern.

The development of DDI as portable metadata standard allows files to be marked up and shared.

In Madison, Wisconsin the CANDDI working group was formed to develop a set of tags appropriate for the Canadian environment.

Statistics Canada and particularly the University of Guelph have been marking up files and discussing problems that arise.

Nesstar may be a suitable commercial solution.
Data Extraction and Analysis System for Ontario

• The DINO Group prepared a discussion paper “Common Metadata Standards and A Centralized Web-based Data Extraction and Analysis system for Ontario

• This paper was presented to the OCUL Directors at their most recent meeting and was positively received. Potential sources of funding are now being investigated

• Thus some province wide data delivery system for Ontario universities seems a possibility
Geospatial Data - Future

- The OCUL Map Group has similar challenges to those facing the data group
- Many of the map libraries acquire the same geospatial data via consortial arrangements
- Each institution is storing, describing and making files available to users
- Many products are very large e.g. Satellite Street View is 256 gigabytes of compressed data, individual files are about 500mb
- How data are accessed, and the support available, differs widely from library to library
Present situation – geospatial data

• Access
  – Some libraries allow selected files to be downloaded directly from a secure site
  – Others require the user to come to the library to sign the licence agreement and pick up the data
  – Sometimes data will be sent by e-mail or via a ftp site

• Support
  – Some have a full time GIS librarian
  – A few have other full time staff
  – Some have little staff support and a librarian with a range of responsibilities
Challenges for Geospatial data delivery

• There are a number of different metadata standards in use such as FGDC
• Much of the data are acquired in software dependent formats e.g. shp, dwg, MrSid
• ESRI with its line of Arc* software dominates the university sector in Canada
• However software preferences vary from discipline to discipline; architects tend to use AutoDesk, engineers may use PCI Geomatics, ERDAS etc. with remote sensing data.
• Open source software is available and the Open Geospatial Consortium is a very active group
Delivery of Geospatial Data

- Software and technology exist to deliver maps over the web to users
- Many approach allows the user to select layers of information and then print or get a pdf of the result
- Some others allow the user to delineate the area and layers of interest and download the underlaying (vector) data. (e.g. www.geographynetwork.ca)
- These options provide an idea of how geospatial data delivery could develop in the Ontario universities
- At the moment the resources are not available to pursue these directions
The Future

• As discussions proceed about the delivery of numeric data in Ontario it is possible that geospatial data will also get considered.

• The future looks exciting in Ontario for data and librarians look forward to the proposal for central storage and delivery becoming a reality.

• It is my hope that the data we have will truly become “Data in a world of networked knowledge” to quote our conference theme.