Integration of "Traditional" Data Centre Services With GIS Technology

by Jeffery Moon[1]
Documents Reference/Data Centre Librarian
Social Science Data Centre, Queen's University

Introduction
Recent cooperation between Queen's University's Documents Unit[2] and the Queen's/IBM Geographic Information Systems (GIS) Lab has resulted in the creation of a GIS satellite site in the Documents Unit. This GIS service point is housed in the Social Science Data Centre and consists of one network-connected IBM PC with related hardware and software.

The Data Centre, Map Library, and GIS Lab are working together through the new satellite site to meet the basic mapping and data integration needs of students and researchers.

An initial component of this project has been the testing and application of Statistics Canada's new graphing/mapping package, E-STAT. E-STAT combines access to CANSIM time series and 1986 Canadian Census data on CD-ROM, and permits these data to be mapped or graphed using straightforward, menu/mouse-driven software. Other initiatives include the installation of OS/2, Windows, and other GIS software, and increased network access to data/map resources.

Geographic Information Systems (GIS) are computer systems that store and link nongraphic attributes or geographically referenced data with graphic map features to allow a wide range of information processing and display operations, as well as map production, analysis, and modelling. [3]

Readers will soon discover that the initiatives described in this paper do not yet meet fully the rigours of this definition. Instead, this paper will describe a cooperative venture between the Queen's University Documents Unit and the Queen's/IBM Geographic Information Systems Lab which, while not yet providing "true" GIS, promises improved service to our users. Specifically, the following will be covered:

Background information on the Data Centre and Queen's/IBM GIS Lab

Integration of services in the Queen's/IBM GIS Satellite Site

Future Developments

Background
Documents Unit - Social Science Data Centre: History

The Documents Unit is part of the Queen's Library System, and consists of the:

Government Documents Library

Map and Air Photos Library

:Social Science Data Centre

The Data Centre was started in 1982 in response to a growing need to consolidate the management of machine-readable data. Its mandate was to act as a clearinghouse for acquiring, archiving, and accessing these data. Initially, support was provided by library staff and one contract programmer. A Data Librarian was hired in 1987 to work half-time in the Data Centre and half-time as a Documents Reference Librarian.

Holdings

Holdings include Canadian Census data, survey data from Statistics Canada and other sources, and time-series data (CANSIM, IFS, CITIBASE). Until recently, the only mapping data held by the Data Centre have been the "CARTLIB" files produced by Statistics Canada. The Data Centre also has several full-text sources on CD-ROM (e.g. Auditor General Annual Reports, CRTC Decisions).

Service
The Data Centre provides a wide range of services including:

:Reference

:Data access/acquisition

:Computing/statistical assistance

:Data conversion/manipulation

Before the GIS initiative, the Data Centre had no public computing facilities. Only one computer was available, for use by the Data Centre Librarian.

Queen's/IBM Geographic Information Systems Laboratory: History

The Queen's/IBM GIS Lab opened its doors in November, 1990. Its mandate is to provide service and support innovation in: teaching (both high school and university), research, and community GIS needs.

Hardware/Software/Networks

Supported in large part by donations from IBM Canada Ltd., the GIS Lab has been equipped as follows:

Computers: 10 IBM PC PS/2 computers

Workstations: 1 RISC 6000 Workstation

General Hardware: a variety of printers, plotters, digitisers

Network: Token Ring LAN and connected to Campus Ethernet Backbone

Operating System: IBM OS/2 V2.0: chosen to enhance/permit - networking, distributed access to ARC/INFO, multitasking & ability to run DOS/ WINDOWS

Software: Arc/Info, Geo/SQL, AutoCAD, Atlas*Graphics, SPANS

Existing satellites: Biology, Education, Health Sciences, Geology and now the Documents Unit

Service

To date, the GIS Lab's major focus has been on its service, teaching, and community involvement mandates. Within Queen's, the primary focus has been on "high-end" mapping (large formats, pen-plotters...) and teaching the theory and practice of GIS. As well, in conjunction with the Faculty of Education, the Lab is currently involved in a two-year project to develop a GIS curriculum for Ontario's Secondary Schools.

In the community, the Lab has been heavily involved in local mapping initiatives and civic addressing. Both of these activities are related to planning for the delivery of emergency services via the "911" system.

Research uses are growing as the user community becomes aware of the flexibility of GIS in non-geographic applications (e.g. life sciences). The lab is working on several projects with health-care planning themes.

Integration:

One of the "problems" faced by the GIS Lab has been its success. As users have become aware of the potential of the mapping services available in the GIS Lab, the Lab has experienced increased demand, including rapid growth in requests for "low-end" mapping such as base maps. Satellite sites created in the first "phase" of the GIS project were neither mandated nor equipped to handle these questions.

To meet this demand, the second "phase" of the project designated the Documents Unit as a satellite site with an eye to providing basic mapping and data integration facilities.

There was some debate as to whether the GIS satellite site should be housed in the Map Library instead of the Data Centre. Given their close physical proximity, and staff's willingness to work together in servicing the the satellite site, the Data Centre was chosen to take advantage of:

:ready connection to the Ethernet Backbone.

:existing hardware, software, and data resources

:expertise with machine-readable data

Hardware/Software/Networks

The hardware donated by IBM to equip this satellite site included:.

:IBM PS/2 Model 57 SX (386, 160 Mb hard disk, 4 Mb RAM)

:IBM CD-ROM player

:IBM Ethernet Card

In addition, the GIS Lab has provided the use of an HP PaintJet (colour) printer.

The Data Centre is connected to the campus Ethernet backbone, which supports FTP and CUTCP. Given this, the potential exists for extensive sharing of files/software between the Queen's/IBM GIS Lab proper and the GIS satellite sites.

As for software and data, the Queen's/IBM GIS Lab has arranged a six-month trial of Statistics Canada's new mapping/graphing package, called E-STAT. A parallel trial is scheduled at the Education satellite site.

E-STAT by Statistics Canada

E-STAT, which had its beginnings in "Telichart", provides user-friendly access to CANSIM ("Canadian Socio-economic Information Management System"), and 1986 Census data on CD-ROM. The CANSIM CD-ROM contains over 175,000 series selected from the Statistics

Canada "Mainbase" of over 500,000 time series. Census data are drawn from the 1986 Census Profiles CD-ROM and exclude all but the Census Subdivision (CSD) Profiles.

E-STAT provides basic graphing facilities for either CANSIM or Census data. A variety of graph types and graph-enhancement options are available. Basically, this portion of the E-STAT package builds on Statistics Canada's CANSIM CD-ROM software, with the addition of enhanced graphing capabilities.

As for mapping, E-STAT provides "boundary files" for:

Canada by province

Canada by Census Division

Province by Census Division

Sub-Provincial Region by Census Subdivision

Only Census data is mappable. Hewlett-Packard InkJet and PaintJet printers are supported.

The mapping portion of E-STAT is designed to let users produce high-quality colour maps with relative ease. Users select desired data (CANSIM or 1986 Census), desired geography (as listed above) and the program automatically generates a high-resolution colour map on the computer monitor.

Retrieved data may be manipulated to create new variables (useful in creating "percent" from "count" variables). All basic arithmetic operations are possible. Maps and graphs can be saved to disk. Map display options include:

Zooming in/out

Panning (shifting the position of the map)

Place name labels/Data value labels (user-selected and automatic)

Changing ranges for data categories

Changing titles

Adding text stringss.

Service
Before the arrival of the GIS-related hardware, the Data Centre had no public computing facilities. Service on ALL fronts has been vastly improved by the availability of the GIS computer, which is not limited to GIS functions.

Access to the GIS computer is scheduled using a sign-up sheet blocked in half-hour intervals. Users are free to schedule time on the computer for any GIS or Data Centre-related function. A range of numeric/text and full-text databases are available on CD-ROM and can be accessed using this machine

From a reference/trouble-shooting perspective, the second computer permits the Data Centre Librarian and a user to be connected to the mainframe simultaneously, greatly enhancing service.

Future Developments
There are several developments underway, and several in the planning stages. Developments are described below, under the Categories:

Software/Data,

Hardware/Networks, and

Service

Software/Data
Software
As new software becomes available in the GIS Lab, satellite sites will be in a position to "upgrade" their capabilities.

For example, OS/2 version 2.0 is scheduled to be installed in the GIS Lab, and will be installed in the Documents Unit satellite site. OS/2 will enable users to multi-task, running different sessions concurrently (e.g. manipulating data in one window and mapping in another, running DOS and Windows applications concurrently).

To provide more flexibility in dealing with user-defined data and mapping needs, several mapping packages are being considered. Primary among these are MAPmaker, provided by Strategic Mapping, and its "twin" MAPviewer from Golden Software (Strategic Mapping bought MAPviewer from Golden Software and renamed it MAPmaker). Both packages run under Windows.

Data
The Map Library will soon be acquiring the "Digital Map of the World", the product of a long-term, multi-nation project. Scaled at 1:1,000,000 this CD-ROM based product will permit direct, user-friendly access to basemaps from around the globe. Access will be mediated through the Documents Unit GIS satellite site.

On a more local scale, digitised maps of "high-demand" areas (e.g. the city of Kingston) are available in the GIS

Lab and will be made available through the satellite site. Once appropriate software is obtained, it will be possible to integrate these maps and existing data sources (e.g. Enumeration Area and Census Tract data from the 1986 Census).

Users can import their own data into E-STAT, to be mapped using the existing array of map coordinate files. This could prove useful, to produce custom maps for geographic areas included with E-STAT, but more flexibility is needed to meet broader user demand. E-STAT would be greatly improved if user-defined map coordinate files could be imported.

1991 Canadian Census data on CD-ROM will be accessible using E-STAT. The availability data for two census years will enhance E-STAT's usefulness for students and researchers. In addition, Statistics Canada is planning to include other social and environmental variables in future releases of E-STAT.

Hardware/Networks

Hardware

Hardware upgrades are already planned in response to anticipated data/software acquisitions. To accommodate OS/2, the GIS microcomputer will need an additional 4-6 Mb of RAM, bringing it to a total of 8-10 Mb. The "Digital Map of the World" requires a math co-processor board. Other hardware upgrades may become necessary as the "computer processing" demands of software and data grow.

On a larger scale, the GIS Lab is using an IBM 550 RISC Workstation as a server, providing access to GIS software (Arc/Info) over the ethernet backbone. Additional mainframe (VM/ESA) disk storage (DASD) is being obtained as well. All of these resources will be available to authorised users, including the GIS satellite sites.

In addition, certain GIS satellite sites are expanding their hardware base to include "X-stations" (diskless workstations) to provide broader access to their facilities at a reasonable cost.

Networks

The Queen's/IBM GIS Lab and all satellite sites will exploit the advantages of network connectivity to the fullest. Common datasets (numeric/text or spatial) will be made available on the network, as will various software packages (e.g. Arc/Info).

On a smaller scale, the Data Centre itself could be networked using "mini-network" packages that permit two computers to seamlessly "share" data and software. This could prove useful, given that the "old" computer is near capacity both in terms of hard disk space and expandability.

Service

The real "test" of the service component of the Documents Unit GIS satellite will come in September when students return. Until then, only limited opportunities are available for evaluating the most effective way(s) of providing service. In anticipation of demand, however, plans have been discussed and groundwork laid.

By their nature, GIS and mapping require printed output. The costs of operating an HP PaintJet are considerable (colour toner - $39.00 Cdn, paper - $24.00/200 sheets). As such, a "fee-for-service" policy has been adopted, with graph or map output costing $1.00 per page. A separate account fund was established along the lines of the existing photocopier system. Determining an appropriate price, though not difficult, did take into consideration the following:

Keep accounting simple, limit need for "change"

Don't make printing so inexpensive that users will print indiscriminately

Don't make printing so expensive that users will monopolise time on the computer to get their map/graph "just right"

Price print output to cover costs

The dual nature of GIS, combining maps and data, makes cooperation between the Map Library and the Data Centre essential. Map Library staff have had to learn those "GIS" applications available at the satellite site and the Data Centre Librarian has had to learn more about providing "map" reference service.

On a related note, the prospects for increased staffing in the Data Centre are unclear. Given current financial restraint, creative means may be required to provide necessary staffing. As GIS and other data-related initiatives develop, it is hoped that a data technician position will be created.

Discussion/Conclusions

The real benefit of having a Documents Unit GIS Satellite stems from the diverse range of resources and expertise this GIS site can draw upon. Combined with the Documents Unit's established user base, this location provides an ideal service point for emerging *and* merging technologies.

Given the early stage of this initiative, prediction is difficult. To date, use of the satellite site in the Documents Unit has been limited to several demonstrations

and practice sessions using E-STAT.

E-STAT's strength is its ease of use. A relative neophyte can produce a colour, CSD-level map in under five minutes. E-STAT was introduced to groups of "enriched" high school students who were given the opportunity to "test drive" the system. The students found the program relatively easy to use, after a basic introduction. For users familiar with other "geographic" software applications (e.g. Atlas*Graphics), E-STAT should present no major problems.

E-STAT's weakness lies in its inflexibility. While users can import data files using the "data manipulation" facilities, user-defined map coordinate files cannot be used. File specifications for map files are not provided. This "closed box" design does not permit E-STAT to be used to its full potential.

Statistics Canada is conducting a survey to solicit suggestions for improving E-STAT. A "new and improved" E-STAT will be released, possibly by the Fall of 1992, based on these suggestions.

Once the academic term starts in September, demand is expected to increase. The availability of a sign-up sheet, and the user-friendly nature of the present software, should assist Data Centre/Map Library staff manage this demand. Key to success will be the continued cooperation among the Data Centre, Map Library, and GIS Lab, working together to meet users' needs. With this cooperation, the Documents Unit's GIS Satellite will be in a position to provide users with access to new "GIS" hardware and/or software as demand warrants and resources permit.

1. Paper presented at the IASSIST Conference 1992, Madison, Wisconsin:.Jeffrey Moon: Documents Reference/Data Centre Librarian, Social Science Data Centre, Queen's University, Kingston, Ontario, Canada K7L 5C4 MOONJ@qucdn.queensu.ca

2. The Documents Unit is administered by the Queen's Library System and is comprised of the Government Documents Library, the Map and Air Photos Library and the Social Science Data Centre.

3. Geographic Information Systems - A Guide to the Technology,John C. Antenucci, Kay Brown, Peter L. Croswell, Michael J. Kevany,Van Nostrand Reinhold, New York, 1991.