Data Intermediation and Beyond: How the Web Modifies the Dissemination of GIS Information

Yang Liang Chua
Sidney Wong

Abstract

As the prominent telecommunication medium, the Internet has four special features making it especially useful to improve PPGIS: low costs of entry, efficient data transfer, interactivity and connectivity. In developing InfoResources West Philadelphia project on an Internet platform, the authors experienced several developments that will shape the future direction of PPGIS. First, the Internet allows the development of a cluster of web sites each accommodating the special needs and technical capacity of users. Second, the interactivity technology forces providers to choose dynamic or static mode of data delivery and such choice greatly determines who will be left out. Third, the Internet greatly expands the data scope. It allows the inclusion of non-GIS data and qualitative information in a single web site, provide an efficient channel for data providers to share and present information, and promote users to develop their community information web sites. Fourth, as the Internet facilitates data collaboration among users and providers, it strengthens the community support and expands utilization. In turn, it promotes participation and highlights the social value of PPGIS in community development.

Using the Internet, however, cannot resolve all the issues in PPGIS because a technological fix never addresses social and structural issues that impede the general users in participating in community development and PPGIS. The Internet will not make a resident use GIS nor help the providers understand the data needs of users. Advanced web applications may create additional barriers against the general public while technology cannot replace trust and committed partnership to build community support for PPGIS. The Internet is extremely fluid. In particular, the information quality and the conventional relationships between users and providers will change because the emerging community information networks will be increasingly supported by indigenous web pages.

Introduction

The advancement of the high-powered microcomputing hardware and the lowering of the costs of desktop geographical information systems (GIS) software have popularized neighborhood indicator programs but achieved little to improve the general public to participate in community-based GIS projects. Public Participation GIS (PPGIS) is limited to small groups of sophisticated users. As such, GIS information providers continue to function as an intermediary between the technology and the general users (Sawicki and Craig 1996). They provide a range of services: data collection, customized data analysis, technology training, and processing raw data into more palatable format. Without intermediary support, most community users cannot make use of GIS information to devise action programs, policy advice, and policy analysis.

However not all community information seekers use PPGOIS services and there is a huge difference in technical capacity among the PPGIS users. If this gap is not widening, it is at least persistent even though technology has made it easier to organize and analyze the vast volume of information and data. There are several reasons for this gap. First, the costs of education, training, and capital investment are much greater than what had been anticipated in the early stage of PPGIS. Second, insufficient resources have not produced enough technically competent staff at the local level. Third, as most community users use GIS sporadically, they have little incentives to spend time and resources in GIS training. Finally, the "have-nots" (those who do not master the GIS skills and lack of support) are lagging behind in the age of the Digital Divide.

Will the Internet, the most prominent telecommunication medium, level the plane and close this gap? This paper argues that the Internet is greatly enhancing the dissemination of GIS information. It takes care of the needs specific to users of different capacity. It opens new avenues to data democracy because community users can create web pages and become information providers. However the Internet is not sufficient to popularize PPGIS and raise the GIS skills of every user to the same level.

The Potential of the Internet

Few years ago, only a fraction of the PPGIS providers used a web page (Sawaki and Peterman 1998). Today it is almost unimaginable that they do not set up one. The Internet has four appealing features making it suitable for PPGIS: a) lowering of cost of entry, b) efficient data transfer, c) interactivity, and d) connectivity. Each feature will be discussed below.

First, the cost of hosting data in a server has come down significantly. For example, some commercial vendors charge as low as five dollars a month to host a web site while domain names cost as low as fifteen dollars a year. The storage capacity of these services is also expanding considerably. If a GIS provider is willing to support the advertisers' banners, it can obtain these services for free. At the user end, the cost of access to a commercial Internet Service Provider (ISP) has become affordable while the access speed has vastly increased to the point that the telephone network infrastructure fails to accommodate. [1] Also users can obtain free access from community networks (widely available in Canada and England) or commercial ISP that requires user to browse advertisement streams. The cost reduction means increasing number of people can use the Internet either at workplace, home or such access points as public libraries, neighborhood computing centers or cyber cafés.[2] The expansion of Internet utilization propels an exponential increase in web sites and, in turn, encourages more people use the Internet. As a result, PPGIS providers can use the Internet to distribute GIS data to a large number of users at any location.

---

[1] The broadband and wireless technologies promise an even faster access speed suitable for instant transfer of video and movie but their cost of entry is still beyond most users' affordability.

[2] Projected from the Department of Commerce's data, more than 60 percent of American homes have Internet access in 2002.

Second, the Internet is extremely efficient for data provision and maintenance. As a communication means, the Internet rivals telephone, fax, and printed materials because of its speed and ability to transmit large volume of information. As a data transfer channel, it can carry text, data, charts, tables, maps, graphs, audios and videos. Once the data are on the server, multiple users can log on to view and retrieve them.[3] Providers can change data offline, upload them to the web, and instantaneously complete the update. These changes show on the user's computer screen even during their visit. This instant data update is far more superior than the paper version kept at public libraries where copies of the updated information must first be printed, shipped to these institutions, and filed on the shelf. It is also more efficient than delivering data by computer diskette. More importantly, the Internet transfers data from one source to multiple points at real time. The data is available 24/7 (twenty four hours, seven days a week) in the web except during downtime. Users can obtain information at their convenience and are not constrained by the operating hours of the data providers or such institutions as the public library.

Third, the Internet encompasses a variety of interactive technologies that enable users to conduct query, select data for retrieval, analyze data, create graphs and maps, plot routes, and generate three-dimensional images. The advent of such technologies as Active Server Page (ASP), Common Gateway Interface (CGI), and Internet Map Server (IMS) has encouraged the proliferation of dynamic database manipulation and mapping functions on the web. The IMS software allows users to perform basic GIS functions without having the application software. Users can customize their data analysis and presentation that were once performed by experts. Ideally, the interactivity technology will reduce the community users' reliance on data intermediation, thus improving data democracy. In reality, this developing technology still requires PPGIS providers substantial amount of expenses in hardware and software. The users also need basic technical training and foundational knowledge in computing, statistics, and mapping.

Finally, as a web-based information system can connect with others seamlessly in the Internet, a web site can indefinitely expand its information storage simply by inserting links to external web sites. In this way, PPGIS providers and non-GIS information providers can work together to disseminate whatever information relevant to neighborhood-based planning and program development. This cooperation promotes competition among data providers and forces them to develop their niches. It also adds context provision to the original role of content provision. This apparent unorganized and redundant system actually promotes data democracy. The quality of data will improve because cross-referencing of web pages allows users to judge one provider against the other. In the Internet, no single provider or information clearinghouse can dominate data provision. As the hierarchical relationship between providers and users is breaking down, users can empower themselves by posting their information in the expanded neighborhood information delivery environment.

The dynamic and ever-changing Internet shows great potential in delivering both GIS and non-GIS data. However can this potential materialized in practice? In the following

[3] The number is limited by the server's capacity to handle multiple requests.

sections, we will examine a web-based PPGIS project (the InfoResources West Philadelphia) to illustrate how the Internet shapes the project development.

About InfoResources West Philadelphia

West Philadelphia is a predominantly African American community suffering from typical inner city distresses. Between 1950 and 2000, its population dropped from 330,000 to 220,000. In 1997, the Center for Community Partnership (CCP) at the University of Pennsylvania (PENN) established a Community Outreach Partnerships Center (COPC). Community leaders expressed a need of creating a neighborhood-based data archive for community planning. In late 1998, CCP obtained a grant from the W.K. Kellogg Foundation and asked the university's Department of City and Regional Planning to develop a pilot project called the "Philadelphia Data Consortium" to deliver community information to community organizations, residents, high school and university students, and researchers.

The project's original mission was threefold: a) generate, collect and deliver neighborhood and GIS data, b) develop an on-line and interactive data retrieval system, and c) perform as a centralized data clearinghouse for Philadelphia's public data. In the first two years, the project was busy collecting data, establishing facilities, and developing collaboration among community partners and university institutions (CCP, COPC, the Cartographic Modeling Lab, Wharton's GIS Laboratory, and later the library). In Spring 2000, Sidney Wong came to Penn and started directing the project. He increased fourfold the staff, scaled back the project, and renamed the project to InfoResources West Philadelphia. The team focused on using the Internet to deliver data. It designed and redesigned the web site and asked for public comments. Eventually in July 2001 it launched the web site for public use.

The fundamental philosophy of InfoR is to promote data democracy and ensure that data belong to the people. InfoR is dedicated to deliver relevant and user-friendly information to users. As such it developed several goals that are slightly different from the original mission:

a. Make the data and information freely available to the general public,
b. Provide useful data for community groups in their grant writing and project development efforts,
c. Translate data into formats that could be digested by community groups,
d. Develop web portals and data sharing projects to connect users to other data providers,
e. Assist community groups and residents to post community information and data in the Internet, and
f. Provide support to researchers and students of all levels in projects concerning West Philadelphia.

The InfoR team studied other community information networks (CIN) and adapted components of the web pages of Cleveland's Neighborhood Link, East St. Louis's EGRET, St. Louis CIN and Seattle's Department of Neighborhoods. It furnishes users with maps, socioeconomic data tables, newspaper articles, and downloadable resource

guides. It arranges its information by such themes as people, crime, poverty, education, health, business and housing. It has a query function so that users can keyword search the web site.

InfoR develops a portal page categorizing over a hundred web pages relating to West Philadelphia and community development. In this way, it practically creates an information archive with an ability to expand. It encourages different information users and providers to share data through the InfoR web site or communicate through its listserve. It allows the public to access a repository of planning studies generated by students, researchers and professors. To add on to this repository, planning students at the University of Pennsylvania are posting neighborhood-planning reports in West Philadelphia from their studio courses while InfoR serves as a data resource for their projects.

Serving diverse groups of users in a community with a quarter million people, the InfoR project is incredibly complex. The site is far from being finished and will be updated when additional data become available. The team will continue to modify its web page design to make it more user-friendly and easier to navigate. It continues to expand its links and encourages individuals and organization to contribute data and information.

Lessons from InfoResources

During the preparation of the InfoR web site, the team had to modify the original project mission because many issues and problems arose. It is not easy to impose a structure upon a complex circumstance where technology is fluid and community interests are diverse. Considering resource constraints, priority of data needs, the lessons from other CINs, and the changing Internet environment, the project team went through a rigorous process of design, pilot scheme, testing, and redesign. It pays particular attention to the opinions of prospective users. Despite these efforts, many issues remain unresolved.

The team learns that it is important to make InfoR flexible and adaptive. One means is to divide the project development into stages. In this way, the team can handle a set of immediate issues relating to the priorities of the project goals and objectives. Some issues should be handled when they are better understood, the cost of entry of preferred technologies is affordable, or the practice of other CINs provides better solutions. The team, therefore, decided that in the short term, InfoR should focus on delivering information on a non-interactive web site, developing a web portal to connect users to other information providers, and including non-GIS information. In the medium term, InfoR will improve data services, develop information sharing collaborative, and assist community groups and schools to post their data in InfoR. In the long term, InfoR will upgrade its technology, adopt interactive GIS (if appropriate), and make the project jointly or solely owned by the community.

The following sections examine four major short-term issues that the InfoR team confronted in using the Internet to deliver community information. These issues are:

- ?? Who are our users?
- ?? What is the appropriate data delivery mode (static or dynamic)?

- ?? What sort of data should be available?
- ?? How are the website and database maintained and expanded?

Serving and Understanding Users

To serve the users better, a PPGIS provider needs to know its users. In other words, it has to understand the users' needs and bring its services close to a specific group of users (Homburg 1994). In doing so, it will develop a niche and makes its service irreplaceable (White 1998).

InfoR targets its services to many groups of users including staff of community-based, faith-based and public-sector organizations, and individuals like students, researchers, faculty and the general public. Every group has their own characteristics. For example, students and faculty who are more technologically savvy and having better support are more inclined to use GIS data. The general public and resident communities at the other end of the technology and support spectrum are less capable or willing to use the InfoR service. Organizational users fall somewhere between these two end and is likely to use the data for specific project.

A web site seems to be an ideal platform for multiple user groups because of its ability to transfer data to a large number of recipients. However in designing the web site, the InfoR team found out the needs of the resident communities had to be met first. This group has slower Internet connection, outmoded computers, smaller screen size, less accessible to technical support, and little knowledge in mapping and socio-economic analysis. Therefore, InfoR emphasizes easy navigational structure, minimum number of images, and digestible tables in designing the web site to ensure reliable connection to all other user groups. In the future, InfoR will add modules targeting to sophistical users to this common platform. In essence, a PPGIS web site can be a cluster of web sites and each gears toward a different user group.

This step has not resolved the intrinsic difficulty in understanding each user group. Can the Internet replace mail survey, forum, and focus group to profile users? E-business sites invariably require customers to register and provide information. Such model cannot be effective in PPGIS unless it serves primarily organizational users (e.g. Neighborhood Knowledge Los Angeles and Philadelphia's Neighborhood Information System). The general public may consider it a privacy violation. A registration requirement likely turns off potential users before finding out whether the site suits their information needs. InfoR decided not to track users and leave the site as unrestricted as possible. Rather, it developed a feedback form in the web site to solicit information on data needs. The users however responded to it poorly, consistent with the findings of studies on the effectiveness of on-line survey. As a result, InfoR is still working to measure the users' need.

The Internet has improved significantly the means to deliver data and information to multiple groups. It also allows PPGIS providers to develop a cluster of web pages sites to suit different user groups. However, the web technology does not guarantee better

understanding of users. The lack of understanding of users is particularly acute for PPGIS sites serving multiple groups.

Data Delivery Mode

"Data delivery mode" is the method of conveying the information on the website. It can be static or dynamic. Under the static mode, PPGIS providers anticipate the types of data that users may require and prepare them beforehand. They then upload a limited amount of information onto the web site. In contrast, the dynamic mode utilizes such interactivity technologies as ASP, CGI, and IMS. In this case, users can query a database, create customize maps, or conduct data analysis. The retrieved information is dynamic (i.e. subject to alternation) and they are prepared in real-time in accordance to the user's specification.

The InfoR has a CGI query functionality to allow users to search index files to locate the information they need. It does not have IMS interactive mapping function because of the following considerations. First, and most importantly, the team is not sure how many users are prepared to conduct GIS analysis because even seasoned Internet users do not have the specialized knowledge in mapping and small-scale geography. As InfoR intends to serve the low end users, the project team does not anticipate any needs for interactive mapping from this group. Second, users who use telephone-line based Internet connection do not experience real time or even reasonable time transfer of dynamic GIS information. Server maintenance, data uploading or excessive simultaneous users often cause down time in interactive mapping sites. The delay and broken communication with the servers further reduce the use of interactive mapping. Third, given the small number of seasoned GIS users, the cost-effectiveness of interactive mapping is uncertain. In addition such setup requires thousands dollars in IMS licensing (with educational discount), and tens of thousand dollars in the upgrade of software and maintenance of the server, its peripheral and back up facilities.

Adopting minimum technology is applicable to PPGIS serving the general public. It is more straightforward and simple because users need not to figure out complex ways to conduct analysis and develop their own indicators. However such a strategy may be less appropriate in situation when the majority of users can conduct meaningful GIS analysis or the provider is willing to spend resources to train potential users. Interactive web site has definite advantages- users have control over the delivery of information and their needs are not limited by the static information that the site offers. PPGIS providers need to assess the cost effectiveness and balance the needs between specialized and general users.

Scope of Data

As the web allows the dissemination of practically any form of information, should a PPGIS provider limit the site as a retrieval device of GIS data? Or should it include other community information? These questions on the scope of data and form of information services actually define the characteristics of a PPGIS web site.

The InfoR team often receives information requests that do not involve GIS maps. In two needs assessment surveys, potential users indicated much needs in non-GIS information including school environment and performance, mailing listing of nonprofit organizations, details of social service programs, grant writing resources, and planning studies and reports. In both surveys, users gave low priorities to census data, property tax and mortgage information. When they are interested in GIS data, their needs are very specific. For example, a street beautification group requested information on the location of trees, while a community development corporation wanted housing condition of individual buildings. Most time, the requested GIS information cannot be met by available public records or administrative data and requires substantial amount of data collection.

The project team conducted additional focus group meetings and substantiated that the users needs were general in community information but specific in GIS data. In other words, if InfoR delivers GIS data only, it neglects the needs of most of the general users – an option that InfoR cannot take. The solution is to provide a mix of GIS and non-GIS data. The team accepts that InfoR is currently unable to meet all the specific GIS needs of few users. In the long term, when resources are available, the team will work with partners to conduct landuse and building surveys, and to integrate the collected data with other neighborhood indicator database. Currently, InfoR offers data and information in the following categories:

1. People - mostly census information
2. Housing – housing loans information
3. Education – schools performance
4. Business – employment and establishment changes
5. Health – health care facilities and public health indicators
6. Neighborhood Resources - directory of nonprofit facilities, social and community programs
7. Crime and Safety – police report on crime
8. History and News- resources on the history and land use development of West Philadelphia neighborhoods
9. Reports - student works, dissertations and other research and planning reports on West Philadelphia
10. External Links – an organized list of external web sites relevant to community development.

The list is constantly growing and InfoR will add a new “environment” category to house environmental indicators. Also the team realizes that InfoR cannot generate all the required data; rather, it must rely on the work of other data providers. The project team pays special efforts to identify relevant web sites and connect them into the InfoR web site. In many occasions, InfoR collects and translates secondary data into more palatable format. The team organized GIS data by a set of standardized data maps and tables at five levels of geographies (census tract, zip code, neighborhood, school district, and West Philadelphia).

A web site is ideal for putting GIS and non-GIS data together. For example, a neighborhood profile can embed quantitative and qualitative information in one web site where text description, graphs, data tables, GIS maps, images and other multimedia materials integrate. In additional, the connectivity of the Internet allows the integration of information from other web sites. It also greatly expands the data scope. The integration or collaboration of multiple data providers has great potential to deliver community information and should be a new direction for PPGIS providers to consider.

Expansion and Sustainability of Site

Data intermediary websites constantly deal with, both financially and logistically, how to sustain its existence. Unlike a commercial web site where profitability governs survival, PPGIS providers receive support mainly from short-term grants. As there is no way to measure the economic value of information services that do not charge the users, the justification of a PPGIS seems to rely on its social value. While social value has no objective measure that funders usually make funding decision based on positive feedback of community groups. In other words, PPGIS providers must develop a network of community support.

At an early stage, the InfoR realized the importance of establishing its existence in the community. It invited potential users to needs assessment surveys and focus group meetings. The outreach arm of the project team keeps dialogue with user groups and works to develop a user support service. As on-line survey did not generate useful feedback, the team puts more attention to develop data sharing collaborative with community groups. For example, it provides technical support to a faith-based organization's senior citizens' service needs and plan to post the survey result in the InfoR web site. Meanwhile, the team is working with a group of high and middle school teachers to integrate the InfoR data in their social studies curriculum. InfoR also uses broadcasting e-mails to market its information services but majority of the effort to expand service utilization is through face-to-face discussion with community groups. When possible, InfoR serves as in-house data consultant of these groups.

The power of connectivity of the Internet has undermined the authority of any single data-clearing house. As cost of posting data on the web is decreasing, small and single-purpose community web sites are proliferating. InfoR considers this development an opportunity for expanding the user base. Incorporating other information providers into InfoR's web site encourages mutual recognition, "coopetition" and specialization. This mutual support is important to promote one another's services to the community. Groups who currently do not generate information may join this network too. InfoR plans to assist community groups in data collection and analysis. By inviting them to post data in the InfoR web page, the team is working on sharing data and hosting these truly community generated information. In doing that InfoR can cultivate more community support, increase the number of users, and improve the project sustainability.

The social value of a PPGIS site should be reflected in its utilization and its contributions in community development. The Internet allows a web visitor counter to record the number of hits though it needs some interpretation to relate the figures to actual

utilization. On the other hand, the Internet brings multiple information sources together. It greatly expands the ability of PPGIS providers to work with community partners in data collaboration, information sharing, and developing a myriad of community information network. The connectivity of the Internet is shifting data dissemination from a one-directional path toward a network of multi-directional paths. This network is more consistent with data democracy because it promotes great degree of participation and collaboration in community information.

The Challenges of the Internet to PPGIS

The four features of low cost of entry, efficient data transfer, interactivity and connectivity provide great potential for PPGIS moving toward greater degree of data democracy. The discussion so far has shown that the dissemination of GIS data is transforming in the Internet environment. The positive aspect of this changes includes the a superior means of data transfer, ability to develop multiple modules to meet needs of different user groups, , the expansion of information scope, the collaboration among data providers through web portals, and the improving ability of community users to generate information. All these indicate that PPGIS has will be more flexible and open-ended. It will be more user friendly and inclusive of the general public.

However, the technological advancement is not sufficient to close the gap between lay and seasoned users of PPGIS. The Internet cannot address the fundamental causes that resident communities consider GIS irrelevant to their life. All the improved communication and on-line tutorial cannot replace the resources and the teaching users skills in basic computing, data analysis, cartography and mapping. Together with the Digital Divide, these are the barriers against popularizing interactive GIS. From the provider's perspective, the Internet does not improve the ability to understand the users. The InfoR experience shows the difficulty in using on-line survey to gauge users' data needs. The development of user support still relies on conventional means to discuss needs and feedback. The Internet is less capable to develop community support too. Building trust and support is through social interaction that the providers and users have to meet face to face. Furthermore, even though the technology is available, it is difficult to popularize advanced applications as the InfoR experience has shown little pay off to the general public.

The Internet is extremely fluid, so there is development ambivalent to the future of PPGIS. New web sites appear as quickly as old ones die out. New wave of technology may quickly make a good model of PPGIS obsolete. As best practice can easily emulate, information providers may have little incentive to be innovative. It is clear that new generations of high school graduates from certain neighborhoods will be technologically savvy enough to be active in using PPGIS. If the Digital Divide is not improving, attempts to introduce new web technology will create more barriers against the lower-end users. This is a great challenge for PPGIS providers to make the service accessible to this group.

Furthermore, it is unclear what kind of information that new generation of users prefer. Data providers have not been very successful to ascertain data needs now, so how can

they plan for the long-term project development? Currently, some PPGIS still deliver specialized data that serve a specific group of users. Some, like InfoR, chooses to present and link whatever information available and relevant to community development. Is this a better strategy? Will it encourage the proliferation of mediocre data delivery and eventually minimize the technical role of a data intermediation? These are new challenges associated with the Internet that remains unclear to PPGIS.

References

Homburg, Christian (1994). What is this: Thing called 'Closeness to the Customer'? *American Marketing Association*, 4: 11-12.

Sawicki, David S. and Patrice Flynn. (1996). Neighborhood indicators: A review of the literature and an assessment of conceptual and methodological issues. *Journal of the American Planning Association*, 62(2), 165-83.

Sawicki, David, and Randy Peterman. (1998). Understanding the depth and breadth of PPGIS Supply. [On-line]. Available: <http://www.ncgia.ucsb.edu/varenius/ppgis/papers/sawicki.pdf>.

Sawicki, David, and William Craig. (1996). The democratization of data: Bridging the gap for community groups. *Journal of the American Planning Association*, 62(4): 512-23.

White, Martin (1998). The impact of communication networks on access to information. In *Information Policy in the Electronic Age*, edited by M. Grieves. London, Bowker-Saur: 165-190.

Homburg, Christian (1994). What is this: Thing called 'Closeness to the Customer'? *American Marketing Association,* 4: 11-12.

Sawicki, David S. and Patrice Flynn. (1996). Neighborhood indicators: A review of the literature and an assessment of conceptual and methodological issues. *Journal of the American Planning Association*, 62(2), 165-83.

Sawicki, David, and Randy Peterman. (1998). Understanding the depth and breadth of PPGIS Supply. [On-line]. Available: <http://www.ncgia.ucsb.edu/varenius/ppgis/papers/sawicki.pdf>.

Sawicki, David, and William Craig. (1996). The democratization of data: Bridging the gap for community groups. *Journal of the American Planning Association*,62(4): 512-23.

White, Martin (1998). The impact of communication networks on access to information. In *Information Policy in the Electronic Age*, edited by M. Grieves. London, Bowker-Saur: 165-190.