

Building the NSDI at the Base: Establishing Best Sharing and Coordination Practices among Local Governments

PRELIMINARY Report
July 2003

Dr. Francis Harvey (PI)
Department of Geography
University of Minnesota
fharvey@geog.umn.edu

Dr. David Tulloch (Co-PI)
Department of Landscape Architecture
Rutgers, The State University of New Jersey
dtulloch@crssa.rutgers.edu

Background

This project began as an extension of two previous projects. Both prior research projects point to the urgent need to better understand and support the informal sharing practices that dominate in local governments. The project *Potentials and Pitfalls of the NSDI for Local Governments*, conducted with UCGIS/FGDC support point to two problems facing the development of the NSDI: 1) lack of awareness and 2) the important role of informal sharing and coordination. The other project, *Exploratory Studies of the NSGIC/FGDC Framework Survey: Looking at the State of the Nation*, pointed to the need to improve data sharing and better understand the ways in which sharing practices occur. The research proposed looks specifically at informal and formal sharing of geospatial data and coordinate geographic information activities and will provide “best practice” outcomes. The proposed research will result in a better understanding of the linkages between existing legislative, political, and regulatory activities and the roles and potentials, and resistance facing any projects to build geographic information infrastructures at the local level. Findings would be evaluated and distilled to produce a list of best practices that can be used in federal and state education activities to support local government geographic information activities.

The project *Potentials and Pitfalls for Vertical Integration for the NSDI* determined that among local governments there was little awareness of the NSDI and very limited understanding of what and how the NSDI could assist. Interesting to note is that local governments with greater NSDI awareness and previous contact with FGDC and other federal agencies on GIS matters also expressed that they saw little support for their needs in the existing NSDI. Their previous experiences had diminished their interest in participation because of requirements and the costs connected to satisfying standards which bore little relevance to their problems. Informal sharing arrangements were working fine.

Complementary to the problem of relevance an underlying issue identified in this research is the importance of existing political, institutional, professional, and legislative relationships. These relationships largely determine the geographic information activities at local government agencies. Sharing and coordination are to a large extent informal activities that correspond to these relationships. Formal arrangements are taken only at the final stages of establishing sharing or cooperation agreements, basically they manifest themselves at the conclusion to satisfy internal procedures and/or legal requirements. This discovery poses an interesting conundrum for developing local government participation in the NSDI. One of the key aims of this project is deepening our understanding of this issue and formulate data sharing and coordination best-practices for local governments around the country.

Given the importance of informal agreements a primary best practices’ question arises: How do local governments successfully share geospatial data and coordinate geographic information activities? The answers to this question are extremely relevant for improving the involvement of local governments in the NSDI. Knowing the issues, potentials, and problems that local governments face is the key to establishing programs and policy support to address local government needs and develop better opportunities for them to participate in geographic information infrastructures connected to the NSDI.

Methodology

Overview Overview

The methodology is divided into three distinct approaches according to the different survey methods used. To establish best practices, the research project relied on two forms of data acquisition. In-depth interviews were held in 6 areas (listed below). These in-depth interviews provided great detail complemented by more tightly-focused polling of local governments. Overall, 91 individuals were surveyed using a basic set of thirteen questions. Responses were either collected by phone, fax-back, or email. Most polling was conducted in the summer of 2002 whereas the interviews were dispersed through the year. The interviews and survey methodologies are discussed in more detail below.

Interviews

The methodology is divided into three distinct phases. In phase one we examined the experiences of data sharing participants with local sharing and coordination arrangements. Objectives of this phase were the determination of strategies and key issues in demonstration projects that have worked and find out what strategies failed and the problems that were resolved and continue to have impacts.

In phase two, interviews were conducted at a variety of local governments to determine how they share geospatial data and coordinate geographic information activities. These interviews were held with numerous local government agencies. Interviews throughout the entire data sharing network provided substantive insights into the contexts and practices of sharing and coordination *in situ*. The local governments whose GI sharing practices were studied were selected with great care. We included large and small local governments representing a diverse geography and worked to ensure that legislative and political differences at the regional and state level could be identified for comparison.

Primary Interviews

The primary interviews generally involved multiple days of on-site interviews throughout the entire data sharing network. Major interviews involved both researchers. The focused interviews involved a single researcher interviewing only the central data producers. Additional follow-up was then conducted by phone from the researcher's home.

Primary interviews were conducted in the following areas:

- Lexington-Fayette Urban County, KY (January 2002)
- San Diego, CA (March 2002)
- Dane County, WI (June 2002)
- Southeastern Wisconsin (June 2002)

The primary interviews provided unique insights into data sharing and coordination practices. To insure we gained insights into the network during our the short visits, we initially contacted the primary GIS unit of the county or region several months in advance asking their permission to conduct a set of interviews. If they agreed, we also asked for the names and contact information for a variety of information sharing and coordination partners. We then contacted these individuals and/or agencies and arranged appointments during a three day period. We tried to leave time open between interviews for post-interview evaluation and additional interviews that arose during our visit. A week before traveling to the destination we emailed all scheduled interviewees a set of questions we would be asking.

Our first interview was always with our main contact person/agency. We wanted to establish their representation of the data sharing and coordination arrangements and environment in their area and develop an overview of legal, administrative, financial, political and statutory issues impacting data sharing and coordination. We then followed with between three and five interviews of data sharing and coordination “partners.” Before concluding all face-to-face interviews and departing, we scheduled a final interview with our main contact to clarify any questions we had and to ask for more specific information that would help us document the areas sharing and coordination practices.

All interviews were taped, after prior consent of interviewees, to assure reliable interpretations and aid later study and comparison of interviews.

Focused Interviews

Although the primary interviews provided a wealth of information, we realized we should also contact areas that had been indicated to us as good models. Lacking the time to pursue comprehensive sets of interviews, we chose to focus on the main contact person/agency with the same questions we raised in the primary questions. The results of these interviews complement the primary interviews by broadening our perspectives.

Two focused interviews were held:

Lake County, IL (August 2002)
Portland, OR (March 2002)

Remarks Regarding the Selection of Interviews

One of the most difficult issues affecting any interview-based research project is the selection of interviewees. Our selections were based on our experiences and our knowledge (through professional contacts and published literature) of the development of local and regional GIS since the 1960s. Two of the locations were chosen based on our previous contacts to people from these sites. Dr. Harvey was familiar with the GIS activities including coordination and sharing issues in LFUCG. Dr. Tulloch conducted research on Wisconsin’s data coordination for his PhD thesis and was very familiar with Dane County and, to a lesser degree, with South-Eastern Wisconsin. San Diego was selected because the county has one of the oldest GIS installations in the country and has had, at times, controversial data sharing. Interviews in these places built on a solid knowledge of their GIS activities. The focused interview locations were chosen following indications from other interviewees that they represent exemplary sharing and coordination examples, in other words established best practices. Lake County is a relatively small, but very rapidly growing county. Portland, OR is renowned nationally for its progressive and successful data sharing and coordination.

Overview of Interviews

In this section we briefly summarize the interviews conducted for this research project.

Lexington-Fayette Urban County, KY (January 2002)

Lexington is a fairly unique example in that it is a merged City-County government. The city/county is relatively isolated. As such, it is particularly well suited for a centralized data supplier and coordinator. Lexington demonstrated how effectively the “hub and spoke” approach can be in a relatively simple setting. While the centralized data supplier carefully controls the flow of data with a pricing policy, public agencies are regularly provided with timely data needed for their specific purposes. Interviews focused on the City-County’s GIS office and users of their data (including the Depts. of Education, Sanitation, Farmland preservation, and the Ky. Dept. of Transportation).

San Diego, CA (March 2002)

Like Lexington, San Diego County is relatively isolated. But the structure within the San Diego County is more complicated with a variety of municipalities and other jurisdictions. San Diego has several different agencies that serve as major data producers. One of the dominant data suppliers is SanGIS, a joint venture of San Diego City and San Diego County (with an additional key role played by a major utility). One of the most important products produced by SanGIS is a county-wide parcel layer. This dataset, of significant value to many users, is licensed on an annual subscription basis.

Another major data producer is SANDAG, a county-wide regional agency, which has produced a variety of planning data. SANDAG uses the property lines from SanGIS as a source of boundaries in their land use/land cover data and other similar data products. SANDAG provides data for free of charge with the exception of SanGIS derived layers when the original SanGIS data is discernable. Data from SanGIS, including parcel boundaries and street centerlines, are not distributed by SANDAG. Data derived from SanGIS data that is not readily discernable or useable in its original form, e.g., the Master Transportation Facilities, however, can be freely shared.

While, SanGIS and SANDAG were the most prominent and in-depth interviews, other interviews included the cities of Carlsbad, Chula Vista, and Escondido, and the local district office of Caltrans.

Dane County, WI (June 2002)

While not a simple hub-and-spoke, Dane County is a relatively isolated situation creating a fairly tight circle of data sharing relationships. It became clear that Madison has a very diverse audience of potential data users including city, county, regional, university, state, federal, private, utility, and non-profit organizations. In Wisconsin, the County is formally designated as a pivotal data producer and coordinator with each county having a Land Information Officer and Land Records Modernization Plan.

Data sharing in Dane County had suffered some recent complications. The County invited a variety of organizations (public and private) to buy into a then-future air photo and photogrammetric project. Without the additional "partners" the project would not have been possible. But to secure the "partners" the County had to treat the data (both images and derived linework) as licensed and limit its distribution. The response to this was wildly varied ranging from outrage to admiration, and has impacted the willingness of others in the area to share their data freely.

Interviews included several different county offices (land information, register of deeds, conservation) as well as the City of Madison's planning office, the University of Wisconsin, and Wisconsin Land Council.

Southeastern Wisconsin (June 2002)

The greater Milwaukee area, 7 counties in total, is within the jurisdiction of the Southeastern Wisconsin Regional Planning Commission (SEWRPC), a major data producer for decades. While SEWRPC has had a long history of sharing a variety of public data freely and for fees, it also administers a joint project with the City of Milwaukee. This project, MCAMLIS, resembles the SanGIS arrangement whereby licensed cadastral and related information was developed and can only be released at a fee under strict licensing agreements. An interview with a GIS librarian at the University of Wisconsin-Milwaukee (UWM) demonstrated some of the implications of this arrangement as it took her nearly five years to reach an agreement with MCAMLIS under which she could actually acquire the data. Despite the history of these policies, other agencies in the area utilize a mixture of policies including a free and open sharing arrangement.

Interviews included SEWRPC, the City of Milwaukee, Waukesha and Ozaukee Counties, UWM library, and Wisconsin Land Council.

Lake County, IL (August 2002)

Dr. Harvey interviewed the GIS manager for Lake County. Lake County was identified by interviewees in South-Eastern Wisconsin as a data sharing leader. A focused interview was arranged with the Lake County GIS Manager. Lake County's GIS reaches back to the 1940s when a WPA project created the first aerial photographic records. Mapping activities were continued into the 1980s when in 1984 the council purchased the first mini-computer for automating parcel records. Because the map department had always seen to have a broad function, even though it became a division of information technology, its roles and responsibilities have always been highly integrative. An idea to create a centralized service center was never considered. Instead, the GIS group has always had a consultative role in collaborating and assisting other departments develop GIS applications. Consultations are also freely offered to municipalities in Lake County. Together with a county-supported GIS user group and training, Lake County's GIS strategy of devolution has sought to create strong end-users who can integrate GIS into their own operations.

Portland, OR (March 2002)

Dr. Tulloch interviewed several organizations in the greater Portland area. Portland Metro is a relatively powerful regional planning agency that gets data from over 20 municipalities, 3 counties and other jurisdictions, and then redistributes these data as a licensed 2-CD data set called the Regional Land Information System (RLIS). Subscription prices vary for government, non-profit, and private interests. While the data is priced fairly low (when compared with comparable places like San Diego) Metro still finds that its map products (often custom) are a much more popular product. Some of the likely data users, like Washington County, have suggested that they rarely use the Metro data because in making it consistent across the region it loses many of the attributes that they find the most useful.

Multi-Modal Surveys

Overview

While the interviews are exceedingly well suited for in-depth understanding of data sharing and coordination practices, they are less well suited to distinguish the uniqueness of the situation from the generality of the issues. For this reason, an important part of our hybrid methodology was conducting surveys of county-level GIS users and developers in Minnesota and New Jersey.

Survey Methods Overview

The surveys were based on the questions asked in interviews, but revised to focus on basic sharing and coordination issues (formal policy, costs, partners, experience, knowledge of metadata and the NSDI). Obviously, the responses lack the detail of the interviews, but respondents desire to share could still be overwhelming. It is worth noting that the ten question interviews would sometimes last over thirty minutes.

Of the 87 surveys distributed to all counties in Minnesota, we received 70 responses, or a response rate of 80%. This extremely high response rate is in large part attributable to the multi-modal survey method and the interest of respondents. Only 13% of respondents were contacted twice. Surveys were completed via phone (21%), fax (39%), and email (40%).

The distribution of surveys in New Jersey followed a somewhat different approach due to the strong geographic and economic relationships to surrounding states. The process began with all New Jersey counties and then municipalities known to us to have produced GIS products. As sharing networks were identified, the interviews reached out to members of those networks (primarily users relying on the local government data) including federal, state and municipal agencies, both those with dedicated GIS staff, and those with personnel who used GIS tangentially. Within New Jersey, the response rate was lower (with only about 17 surveys fully completed), but the surveys were treated as full interviews.

Parallel to these interviews, additional similar interviews were also conducted with about 20 New Jersey non-profit organizations that rely on local and state government geospatial data. These interviews provided additional information and helped guide the identification of known data producers. This helped make the process of identifying likely data producers in New Jersey much easier.

In distinction to the multi-modal survey methods used in Minnesota, all survey data collection occurred through telephone interviews, usually cold-calling to request a 10-15 minute on-the-spot interview with GIS personnel. Interviews were semi-structured, generally following the survey order, interviewees deviated from the survey order to describe specific projects or to explain the relationships between sharing agencies. Some interviews lasted as long as 25-30 minutes.

Interviewees generally were eager to participate, though a few were wary of discussing data policies, or used sub-contractors to develop GIS products and therefore didn't know enough about geo-spatial data to discuss data issues. Those individuals who were willing to be interviewed typically were forthcoming about data sharing policies and interested in the results of the study. Interviews using the same protocol were expanded to the greater Portland, Oregon as part of the follow-up to a visit by Dr. Tulloch.

Remarks Concerning Multi-Modal Survey Methods

On the other hand, some respondents indicated a lack of time to conduct the interview over the phone. For this reason, we introduced a fax-back version of the survey. The protocol was modified by introducing a preliminary question whether the person would like to take the survey over the phone (21% of surveys) or have the survey faxed to them for filling out and faxing back. The fax-back option was quite commonly chosen. Of the 70 surveys returned in Minnesota, 39% were returned by fax. The same procedure was used for surveys sent out by email: 40% of returned surveys were emailed. The question arises whether the results from these different distribution methods are comparable. Based on our review of survey literature, we argue they are, but we need to bear in mind that the give-and-take of phone surveys offered respondents more options to query the interviewer as to the meaning and intents of questions. This discourse was unavoidable, but to limit its impact we instructed interviewers to politely indicate to interviewees that "what they thought the question meant" was how they should interpret all questions. In other words, the greater detail of phone interviews was readily balanced by the specificity of written responses.

Evaluation

Our interpretation of interviews began during the first interview, held in January 2002. At the end of each set of interviews we summarized our overall impression through a "closing discussion." Survey responses were analyzed by project staff and brief summary reports prepared. We use

these summary reports in preparation of the final report and of other research publications. Recordings of interviews were consulted to fill gaps in our memories and notes.

The results of our evaluations appear in next section of this report.

Overview of Results

The research results of this follow-up project focus on developing best practices for local governments. This includes generic best practices and specific practices by type and situation of local government. These best practices would also document strategies that have failed. These results will not only point to the “ideals” of GIS, but also indicate approaches that didn’t work and problems that derailed GIS projects.

These best practices will be invaluable inputs for local government GIS implementation and consortia building, and participation in the NSDI. They will also be useful as insights into local government practices for policy formulation at local, regional, state, and federal levels.

Sharing/Coordination

As previous studies indicate, sharing of geographic information is wide-spread and pervasive at all levels of government. The results from this research underscore this finding. Of all 57 survey respondents, only one respondent from the counties that used GIS was not sure if they share data. All other counties indicated varying degrees of sharing, ranging from 1 to 15 or more sharing partners. The only county responses that clearly indicated no sharing were counties without any GIS software/hardware installed.

Coordination is, however, another issue. Responses to this question took into consideration that we were considering joint activities including data acquisition, processing and dissemination. Clearly, the complexity of organizing these activities can partially explain that over half of the respondents do not coordinate any geographic information activities. That close to half of the respondents coordinate activities, however, suggests that coordination becomes a significant activity, largely with state agencies. Sharing and coordination activities are far less common with other counties and municipalities.

Data Distribution

The most common forms of data distribution, as would be expected, are provision of data in electronic format, either by web (ftp) or email. Almost as frequent is the distribution of data on CD/tape. Very few respondents indicated that they review requests for data, although more often they would request a simple written request.

Most striking is the lack of distinction between different request categories (same jurisdiction, another local government, known government agency, or state agency). Preferred data distribution is the same for all categories. However, different categories face different conditions. Half of the respondents in Minnesota who were involved in data sharing indicated that they applied conditions to data distribution, the other half did not. About 9% do not share data. In New Jersey 13 respondents required an agreement, two did not, and six indicated the question was not applicable.

The conditions for data distribution vary, but a few specific issues emerge from responses in Minnesota and New Jersey. Perhaps the most common condition placed on shared data is “No further distribution.” The reasons for this range from economic protection to control over data, but seem largely to reflect concerns for potential misuse of data. This explains why liability is

also a common condition for data sharing. Although liability release forms are uncommon, these two conditions are often woven together in a disclaimer. Disclaimers can be specific: “data only used for requested purpose & not shared w/others w/o permission” or broader: “data use & acceptance; no guarantee of format; copyright; no resell or redistribution via web; no responsibility for manipulated data.” In all disclaimers, statements about the quality of data are infrequent, whereas statements about the quality limits or lack of known quality are commonplace. Other conditions include requirements that the data source be indicated in all products, either in the form of a citation or reference note. Data privacy also appears as a condition, but is largely affected through data sharing policy and practice, e.g., “no sharing of address data outside of the county.”

Intriguing remains that in spite of increased availability of data over the internet and more formalized data request procedures, informal sharing still plays an important role in some places. For example, in one rural Minnesota county that largely shares data with the state DOT and municipalities within the county, informal sharing predominates. In their words, an “I owe you one” is usually the method of sharing.

Use and Role of Standards

Striking, and highly intriguing, is the clear indication that 40 respondents in Minnesota relied on standards, 19 did not, and 12 indicated standards were not applicable, or possibly used, or simply asked “what are standards?” Examining the detailed remarks to this question, it is plain that standards are understood in a myriad number of ways. For some respondents standards are documentation of metadata: “do our best to document the metadata for each layer and identify the positional accuracy.” Other respondents coordinate standards usage with others: “confer w/ contracting agency, GIS Users Group to use same stds and software compatibility.” Some respondents have even developed their own standards to specifically address their prerogatives: “personal standards higher than state & fed; nothing written in stone.” At the same time this response highlights that standards are always changing. For a variety of reasons, the use and role of standards can change very quickly. A few individuals are frustrated by this and other complexities of using standards as the following verbatim quotes indicate in response to this question: No-“when I'm aware of standards that I can meet” or “standards??? Order of priorities- real time needs first.”

Metadata was the most common standard mentioned by respondents in New Jersey. Those who actively maintain metadata mentioned other standards such as XML, GRASS, ARC, State Plane Coordinates, etc. In New Jersey 48% indicated they used metadata compared to 60% in Minnesota. While CSDGM is certainly known, many respondents use it partially, modify it, or have developed their own formats. For some, but by no means the majority, voluntary compliance with state metadata standards was also important.

Data Sharing Agreements and Policy

Responses to the question about attempts to organize data sharing and/or provide policy for data sharing clearly indicate the tensions local governments face. While data sharing organization and/or policy are common (40 respondents in Minnesota indicated they have some organization or policy for data sharing, 25 do not, and only six indicated the issue is not applicable), it is not equivocally implemented nor resisted. On the extremes, negative responses indicate a strict limitation of sharing: “do not share other than what is required by law.” Positive responses to this question point to vastly different sharing philosophy: “our policy is that we share data freely w/ gov't entities.” Regardless of the actual policy or arrangements respondents overwhelmingly refer to serious concerns and problems facing data sharing arrangements and policies. Problems range from difficult relationships with state and federal agencies, funding,

priorities and tangible deliverables for county activities, desire to coordinate with sharing partners, data sensitivity, data privacy, etc. Few respondents indicated a completely free and open data sharing arrangement or policy. Even those with FTP or Web sites made clear that the sites were only accessible on the intranet, or required a previous agreement.

In New Jersey more can be said about the specifics of data sharing arrangements and policies. New Jersey counties (and others, like Rutgers University and non-profits) tend to use standard New Jersey Department of Environmental Protection (NJDEP) sharing agreements (see Appendix N). Specifically, a common approach is to modify the NJDEP's shrinkwrap agreement included with its freely-distributed data CDs and downloadable data. This reflects the fact that NJDEP is probably the most visible data sharing organization in the state.

Knowledge of NSDI

As previous studies have indicated, the NSDI is still widely unknown. About 46% of the Minnesota respondents indicated that they have heard of the National Spatial Data Infrastructure. 36% had not heard of the NSDI and 10% said it was not applicable. In New Jersey, 61% of the respondents indicated they heard of the NSDI. Only 1% had not heard of the NSDI, but 29% said it was not applicable. While these results by themselves suggest a variability in the knowledge of the NSDI in local governments across the United States, comments made by respondents who had heard of the NSDI in both Minnesota and New Jersey suggests some fundamental misunderstandings of the NSDI and an overall very limited knowledge of the NSDI. One respondent added this comment, "NSDI seems to emphasize sharing of geo-data and promotes/creates stds/guidelines to assist. I believe stds such as SDTS evolved from this." Demonstrating less misunderstanding, but a plain frankness that also underscores the lack of understanding, another respondent said, "heard of it but don't have a whole lot of info about it." If nothing else, pointing out not only the problems of developing the NSDI, but also the problems of data sharing and coordination is this comment: "lectures, consortiums, websites - A wonderful idea that will be difficult to execute."

Risks of Data Sharing

The limited knowledge of the NSDI and misunderstanding of its concepts is all the more troubling considering the concerns with risks of data sharing. Liability is one perspective on the problem. Overall respondents took a far more nuanced perspective by balancing risks with benefits. In Oregon, the extensive use of GIS data in planning activities seemed to open up more risks: "we can't control whether the data is the best for a client's intended use" i.e. earthquake data at wrong scale for use, planners use county data that is incomplete. The agencies that dealt to some degree with private organizations (HMDC, Green Acres) saw specific risks affecting property value such as the distribution of certain data could drive up the price of properties they want to obtain; or the release of sensitive data such as hazardous waste sites might cause problems.

Most respondents in New Jersey and Oregon consider that the issue of risks is not applicable (seven out 21 responses). In Minnesota the picture is far different: 75% of the Minnesota respondents identified some form of risk with data sharing. Only 1% found no risks and 24% considered that the risk issue is not applicable. Responses to this question in Minnesota contained a wide-variety of risks. The most common risk was misuse of data followed by misrepresentation. Specific risks were also mentioned, for example, the misuse of feedlot data. Respondents who also indicated they did not share data, except as legally required, saw the greatest risks as this response suggests, "data is dangerous unless constantly updated."

Evaluation

The surveys and interviews provide rich material for this research project. In our evaluation, we will focus on the interpretive sociological analysis and identify three types of data sharing and coordination that exemplify best practice approaches. Typifications consider many characteristics: organizational, community, political, legal, financial, personalities. These issues could be lumped together as context, but lumping diverse characteristics into context runs the risk of ignoring critical aspects and relationships. Each typification is an ideal which we are as likely to meet as the average citizen. Our typifications focus on patterns and processes which we identified during our research.

Type 1: Hub-and-Spoke or GIS Dictator

Brief Description

The hub-and-spoke model is a centralized model of data sharing or distribution. It is alternatively referred to as the “GIS dictator” model as a reflection of the tendency for the centralized agency or individual to be a controlling factor in the entire data sharing network. The “hub” is generally a major data producer as well as coordinator and sharer. Often times data sharing between two other agencies would go through the central agency. Beyond simply facilitating data sharing, the hubs that we found were often providing additional technical assistance which furthers the reliance that the “spokes” have on that central agency. Generally speaking, the spokes have little choice about whether to work with the hub. These central producers are often one of the most technically sophisticated GIS users in the area.

Advantages

A strong and benevolent “dictator” can work to promote more and better use of GIS and its data. They can use their power as a way of enforcing data and metadata standards. They also can serve as role models to the other data producers and users. They serve as easily identified sources of information (e.g., How good is the data? Just call John. What’s the coding mean? Just call John.) Successful “dictator” have substantial political capital with many groups and individuals which helps in weathering economic downturns and political changes.

Disadvantages

One of the obvious problems can be an abuse of their power. They can control their data very tightly in ways that benefit collaborators but not only disadvantage groups without access, but can preclude them from participation. Unless mandated, they rarely make it their highest priority to help users outside of the data sharing network. The hub can also potentially withhold data to maintain and/or increase their power, and encourage poor metadata and data standards to assure people with questions have to return to the hub.

A major concern, even with a “benevolent dictator,” can be their departure. As strong centralized data producer and distributor can be fairly hard to replace if they were the only one who knew what their coding meant or where the county street files were saved.

Type 2: Federation by Accord

Brief Description

The federation is a more complex network of relationships than the Hub-and-Spoke model. The federation involves a fairly good number of data producers who generally share their data with a number of other data users and producers in their network. While there is often a central data producer or two that cut across the jurisdictions of many, they lack the singularity or strength of the hub. But, in the federation by accord model, they are often allowed to serve as a coordinator and clearinghouse simply as a means to facilitate the work done by the other.

Advantages

A fairly orderly form of chaos. Can afford to lose a major player without ruining the entire model. This model places great emphasis on a philosophy that government is there to serve the people. Over time, this model builds thick networks of connections and relationships that help assure its longevity. More than data sharing, this model promotes exchanges, collaboration, and interactions. It approaches the ideal data sharing environment in many ways.

Disadvantages

Relies on “accord” and relatively equal status among the major participants. There can be significant inconsistencies in format, quality, and presence of data across the jurisdictions of the federation. Personalities are key. Corporatist approaches clash with this model. If a major producer tries to charge exorbitantly for widely needed data, the other members of the network might feel the need to “match” the cost policy or seek recourse. There is a greater “unknown” involved in a voluntary system – will all of the data always be available? Does all of the data exist across the federation?

Type 3: Federation by Mandate

Brief Description

In some circumstances, an agency (or group of agencies) is given special authority in regards to data production and sharing. An example would be a regional planning agency that is designated by the state as the official producer of specific data layers. Their authority may extend to requiring other jurisdictions to submit data to be incorporated into the official dataset. Unlike the hub-and-spoke, this would be a more complex network, in which many of the participants are major data producers.

Advantages

The mandated participation creates a significant opportunity for consistent data across a jurisdictionally complex landscape. The federation by mandate has a high reliability over time, wherein participants can generally count on the availability of the same data under roughly the same terms year after year across a fairly wide area.

Disadvantages

Unlike the Federation by accord, the mandatory participation can sometimes result in hard feelings and abuses of power. An over-reliance upon the lead organization can also lead to poor work on behalf of the “subjects” of the federation. Tensions may lead relatively quickly to

scuttling this approach. This can of course happen with great regularity considering the two or four year terms of most elected officials.

Summary

In conclusion, we offer a synopsis of the major findings of this research project. These findings focus on best practices as evidenced from the project's surveys and interviews. While the majority of the findings lead to duplicatable activities, we also want to underscore the contextuality of most best practices. Without a "cookbook" the application of best practices involves a great deal of learning: learning by reading about other group's experiences, learning by talking at conferences and meetings with other professionals, but mainly learning by doing. It is ultimately the specific circumstances and possibilities unique to each data sharing and coordination activity that define best practices. With this in mind, we also want to highlight five contradictions of best practices that should provoke all kinds of learning

Seven Key Points

Context matters

Different institutions require different responses. There is clearly no one best practice that can be offered. Decisions about whether data sharing should be formal or informal may depend on where you are. No matter how badly you want to give away the data, if you work in "Tammany Hall," you can't. Some isolated systems may be able to rely on a hub and spoke model while some complex multi-jurisdictional landscape may require fairly sophisticated models to ensure that data is available in appropriate and fulfilling ways.

Attitudes vary

When asked whether they charged, some people asked, "Why?" When asked if they shared data, some people asked, "Why?" Some people saw the data as the source of their power. Some people saw giving it away as their source of power. Sharing scares some people. Some people have concerns about the risks involved -- liability is a big concern. (But some URISA folks are suggesting that liability goes up when you charge.)

Charging for data can cost more than you think

(or Be wary of the ripple effect) Sometimes, when you think you are making money off of your data, you are really costing more than you think. Charging for data can have primary, secondary, and tertiary economic effects that you should be aware of before you chose a restrictive access policy.

Bigger is better

Generally speaking, the larger organizations seemed more likely to share, to have developed metadata and to be prepared to participate in a larger SDI.

Where there's metadata, there's data

While not all data come with metadata. We found that folks who kept metadata almost always had lots of data and it generally seemed to be data worth getting. But many agencies admitted to using standards that were not FGDC compliant

Sometimes, it's all about who you know

A number of institutions explicitly admitted to sharing data freely with people they know and trust, while making it difficult for others to gain access. It became clear in some places that almost everybody had studied at the same school, so that even if they weren't classmates they

shared the same favorite faculty. These personal connections seemed to really overcome some other limitations. It can also be about how you treat them. When you are working in a hub and spoke environment, you clearly need to treat the GIS dictator nice. If he/she shuts you out, it can be very hard to get back in.

Sharing is easy, not sharing is hard

Just giving your data away can turn out to be the easier and more affordable route. Copying someone else's shrink-wrap agreement and leaving your data on the web page can be pretty simple. Dealing with lawsuits, chasing down "illegal data launderers", and negotiating iron-clad license agreements can be VERY hard and unrewarding work. Some agencies seem to spend more time and energy dealing with preventing "data theft" than they make in their cost recovery charges.

Five Best Practices' Contradictions

1. Remember, to give is divine, but knowing you might get something back later is pragmatic.

Data sharing is a good thing—as good as mom, apple pie, and the flag. While economic considerations may motivate people to hold on to data or charge for it, most people recognize that data collected with public funds is a public good. Sharing public goods freely is as important to the democratic process as voting. All the same, it is undeniably important to know that when you share data, you (and/or the department) will be recognized. Recognition is an intangible currency of public administration, but one of the highest values for an organization and local politics.

Unexpected ways of receiving are also very important. An agency might pass on new data to a contractor and receive a windfall at the next budget meeting because a council member was able to ensure the development of a new park thanks to the contractor saving money on survey costs.

2. Give the data away, but always make sure people know where it comes from.

Obviously giving data away makes sense—for any number of reasons. Also because no one has yet been able to demonstrate a successful cost-recovery program in the U.S. You also need to be sure that people know where the data came from. If they don't know where it came from they don't know who to contact about updates, changes, possible uses, and suggest improvements. People also don't know who to thank and recognize.

3. Don't charge for data, but make sure people know what it costs.

What's the value of something you get for free? Nothing, claim pundits. That may be so among people who deride the responsibility of government and wish to diminish its role. However, people who know the value of something are generally more appreciative of its value. Agencies that communicate the value of data they share for free can be easily recognized. Of course, determining value is a complicated affair. Be sure to be judicious and conservative in estimates. People value a good value even more.

4. If possible, never charge fees, but make additional services available at a fair cost.

If revenue and cost-recovery are important issues, instead of fees, many agencies have been successful in providing additional services to the no-cost provision of data. This can be done alone by a department, but public-private partnerships can be important way to create some jobs in the community and strengthen local bonds. Remember, that data is just the smallest part

of any GIS. People and analysis are the largest parts and they require resources. You might even want to think about offering training support to give people from other departments and communities a solid start. Once they know what can be done, they will be back many times for more data and help.

5. If you have data, remember, you don't own the room that people are eating in, you're only putting the "food" on the table.

The last time you went to a catered party (or set up the party yourself) do you remember the caterer decorating the room or telling the host/hostess how to greet visitors? Since this is unlikely to ever happen, remember that when you share data, you are only dealing with part of the activities. It may be the most important part, but remember who is using the data and its importance for their activities. Help with the data to make their work a success and the benefits will be bountiful.

Conclusion

Best practices of data sharing and coordination involve many aspects of administrative and political activities. The best practices for any particular locality at any particular time are contingent on a number of factors and characteristics. Of all the practices we have identified in this research, a key practices seems to lie in the approach towards data and colleagues. If data sharing and coordination is just about the data, it will be very difficult at best, and may likely not work at all, or for any length of time. Data sharing and coordination are best understood as part of other activities. Some of these activities require interaction, many others are assisted and promoted by data sharing. Establishing and supporting a social network among colleagues, citizens, and elected officials that supports their interests seems to be critical in all cases.

Appendixes

Appendix 1: Interview Questions

Question Domains

- 1) Data sharing
- 2) Autonomy
- 3) Financing
- 4) Consortia
- 5) Vertical and horizontal integration

Questions

1) What does data sharing mean for you? What activities does it encompass?

2) Who do you share and coordinate data acquisition, processing, and dissemination with? (For a range of partners) For each one:

What data? Whose data is it? How often? Is it a one-way or a two-way relationship? Is it a one-time exchange or a re-occurring activity? Or part of other governmental activities?

In cases of you acquiring data: How did you find out about the potential of sharing this particular data? (they came to you, clearinghouse, advertising, word of mouth, long-term institutional relationship, political relationship)

What is exchanged? Just data? Metadata? Ideas and relationships?

Are conditions stipulated upon the sharing, the use of the data, redistribution? Are they recorded in contracts or through memos of understanding?

What are the monetary costs associated with the sharing?

Do you have to pay to get the data?

Who paid for its creation?

Who pays for the dissemination processes?

Are there any "hidden" costs? How does it effect your bottom-line? What are the consequences in your council or executive board?

What is the motivation for you and the other party/agency? Do think both of you get what you want out of the data sharing?

If this data sharing relationship came to an end, could you work around it? Would you bother?

3) How does data sharing impact your intra- and inter-agency relationships and budgetary negotiations?

What risks do you associate with sharing data?

How does it impact your department's or agency's organizational policy?

Do you, or your colleagues, ever sense a potential to loose your autonomy?

What standards do you deploy or consult?

How common is informal data sharing?

What are the "best practices" for you?

4) What strategies have worked best for your organization to assure your financial solvency?

5) Can you briefly trace the use of GIS in your organization? Please provide a basis to understand the “historical” development of your agency’s use of GIS?

6) Have there been any attempts to establish organized data sharing and/or to provide policy? What about these attempts was successful, what was unsuccessful?

7) Please characterize your relationships with regional, state, and federal agencies in terms of data sharing?

Have you made specific efforts to share data with these agencies?

Have you been involved in sharing data between agencies in your municipality that simultaneously involved regional, state, or federal agencies?

Please describe the open-records laws of your state to your best of your knowledge.

What impacts have they had on your data sharing and coordination policies?

8) Have you heard of the National Spatial Data Infrastructure?

When?

Do you know/use any standards proposed by the NSDI?

Are you aware of their funding programs?

Have you collaborated with FGDC sponsored activities?

Do you use the FGDC/NSDI Clearinghouse either to access data/metadata or to distribute it?

Appendix 2: Survey Questions

Introduction

We are interested in how you share and coordinate geographic information. {some information about what we understand} We'd like to find out from your perspectives on sharing and coordinating, with as much specifics as possible.

1. To start out, can you tell us something about your activities in <county, consortium, etc>?
2. Any exchange with regional, state, or federal government agencies?
3. If someone comes to you and asks for data what would you say? (Use examples and Check for local, regional, state, federal, and public)
4. Follow up with question about specific gov. activity, e.g. when a house is built who do you get the address and other information from?
5. What about the costs of data? Do you pay for data? How do you pay?
6. How long have you been using GIS (both personally and agency)
7. Is there ever additional data that comes with the GI to help explain the GI? (Metadata)
8. What is your background?
9. Have you ever heard of metadata?

10. Have you ever heard of the National Spatial Data Infrastructure?

Thank you very much for your time. If any questions arise later, would you mind us contacting you?

Appendix 3: Data Distribution Agreement (NJDEP)

Agrees to abide by the terms and conditions of the following:

I. Description of Data to be Provided

The data provided herein are distributed subject to the following conditions and restrictions.

Subject Data Layers

For all data contained herein, (NJDEP) makes no representations of any kind, including, but not limited to, the warranties of merchantability or fitness for a particular use, nor are any such warranties to be implied with respect to the digital data layers furnished hereunder. NJDEP assumes no responsibility to maintain them in any manner or form.

II. Terms of Agreement

1. Digital data received from the NJDEP are to be used solely for internal purposes in the conduct of daily affairs.
2. The data are provided, as is, without warranty of any kind and the user is responsible for understanding the accuracy limitations of all digital data layers provided herein, as documented in the accompanying Data Dictionary and Readme files. Any reproduction or manipulation of the above data must ensure that the coordinate reference system remains intact.
3. Digital data received from the NJDEP may not be reproduced or redistributed for use by anyone without first obtaining written permission from the NJDEP. This clause is not intended to restrict distribution of printed mapped information produced from the digital data.
4. Any maps, publications, reports, or other documents produced as a result of this project that utilize NJDEP digital data will credit the NJDEP's Geographic Information System (GIS) as the source of the data with the following credit/disclaimer:

"This (map/publication/report) was developed using New Jersey Department of Environmental Protection Geographic Information System digital data, but this secondary product has not been verified by NJDEP and is not state-authorized."

5. Users shall require any independent contractor, hired to undertake work that will utilize digital data obtained from the NJDEP, to agree not to use, reproduce, or redistribute NJDEP GIS data for any purpose other than the specified contractual work. All copies of NJDEP GIS data utilized by an independent contractor will be required to be returned to the original user at the close of such contractual work.

Users hereby agree to abide by the use and reproduction conditions specified above and agree to hold any independent contractor to the same terms. By using data provided herein, the user acknowledges that terms and conditions have been read and that the user is bound by these criteria.