MAKING METADATA A PART OF YOUR DAILY DIET

Abstract: Metadata requires no introduction, though for the record, we're talking about background information for a dataset, data about data, or the recipe for your great grandmother's apple pie. Just think if your grandmother hadn't written down those ingredients or cooking instruction- it wouldn't taste the same!

It appears that metadata is leaving a bad taste in the mouths of many GIS gourmets, in a manner analogous to garlic. Raw garlic tastes awful, but the professionals espouse garlic's health benefits. Metadata has been the topic among professionals for some time, but now the time has come to find ways to make metadata more palatable to the GIS community so that data developers will voluntarily complete metadata for in-house management of their datasets and for data sharing transactions.

INTRODUCTION

Metadata

When the ancient Mesopotamian people invented beer, they could see that the idea had a lot of potential. Though most people couldn't read at the time, the king's scribes made an effort to write down the recipe for brewing beer. If the king's official brewer were to have died without an apprentice in place, the art of beer making could have been lost, not to mention a few heads. With that in mind, the incentive existed for people to document information at the nascent stages of world civilization.

Though people have documented for centuries, the term "metadata" had only appeared within the past decade.

Move ahead about four millennia and millions of liters of beer, and the word metadata appears possibly for the first time on the planet. According to a source on an internet metadata discussion line, the word "metadata" was first used in the late 1980s. The year was 1988, and NASA was the source, with metadata mentioned in the context of a Directory Interchange Format (DIF) manual. The DIF format is used to create directory entries that describe a group of data. Much like metadata, a DIF allows data users to understand the content of a dataset, using fields that provide the information necessary for users to understand and decide if a particular dataset would be useful for their needs.

Over the course of ten years, metadata would follow many paths, including its evolution along side of the Geographic Information System (GIS). This evolutionary process of metadata has been a salient issue that continues through today as professionals discuss such topics as metadata standards and software tools.

Interest in metadata is moving from concept to implementation.

What appears to be lost in the discussion is metadata implementation. While it is critical to define metadata standards and provide effective software tools to create metadata, it is also necessary to discuss implementation. As it stands now, the best standards and software tools available may not entice data producers to sit down and write metadata. Probably the least attractive responsibility tied to GIS and data development, metadata creation remains a low priority amongst the data developers.   The reasons are many, ranging from a lack of familiarity with metadata to time and resource constraints. This paper will address these issues and explore the approaches used to encourage metadata implementation among organizations that use GIS in the Twin Cities metropolitan areas of St. Paul and Minneapolis, Minnesota.

The ancient scribes and contemporary GIS producers share many of the same goals of retaining information to assure that the value of data is maintained for themselves and all users. The difference is that data producers can be assured that a lack of compliance will not mean getting impaled on a stick.

MetroGIS and Metadata

MetroGIS is a Geographic Information Systems project that is organized to provide data sharing assistance to counties, cities, watershed districts, school districts, state, federal, regional and non-government organizations serving the Twin Cities seven county metropolitan area.

Since its inception in 1995, MetroGIS has navigated a course towards obtaining data and cost sharing agreements between all Twin Cities area counties and the Metropolitan Council. The Metropolitan Council, a regional government organization, facilitated MetroGIS creation, and provided a majority of the funding and impetus for many of the data sharing efforts across the metro area.

Currently, the future success of wide-spread data sharing activities hinges on the development of Data Finder, an internet web site[1] that the MetroGIS project launched in 1998. Data Finder provides the mechanism for the Twin Cities Metropolitan Area data producers and

---

[1] www.datafinder.org

2

D users to display and share information about available digital geo-spatial datasets. The source for dataset information is metadata.

Metadata represent the key to Data Finder's functionality and its survival, which, as mentioned, drives the data sharing activities. Without metadata, GIS organizations lack the capacity to effectively exchange information about their datasets, thus seriously affecting the project's intended goal of data sharing.

MetroGIS' METADATA IMPLEMENTATION STRATEGY

MetroGIS sought to inform data producers and users about metadata, evaluate responses to metadata, and explore ways to encourage metadata compliance.

Recognizing the importance of metadata for the Data Finder project, MetroGIS funded a position to focus specifically on metadata implementation among GIS data producers and users in the metro area. The Metropolitan Council hired a metadata assistant in August, 1998 for a one-year term, and handed him the responsibility of getting the word out about metadata and their benefits. Another duty would be to explore ways to encourage metadata compliance, plus evaluate the strategies employed in these efforts. Written assessments and evaluations of activities would be conducted on a regular basis in an effort to know which strategies proved most effective, or had potential for future success. Also, as part of the metadata project, an agreement was reached with one of the seven metro area counties that would permit the metadata assistant to create a department's GIS metadata for their datasets. This experience did provide some insight into the methods employed to gather information for metadata, and take into account time, costs, and efforts for an outside contractor to create an organization's metadata.

In the end, MetroGIS was able to draw knowledge from this cornucopia of experience and information to establish a foundation for devising strategies that could be tailored to the fit the aspirations and limitations encountered with each organization's response to metadata.

Metadata Implementation Tools

Tools for the implementation effort:
1) Description and background information;
2) Description of Benefits;
3) Data Needs assessment study;
4) DataLogr software;
5) Minnesota Metadata Standard.

The metadata implementation project would involve several stages, or strategies to advance the concept and acceptance of metadata among the targeted organizations in the metro area. Disseminating the concept or idea of metadata would require a substantial arsenal for the first stage of this project. At hand for the first stage of this endeavor were the Minnesota version of the FGDC metadata standard, a metadata software tool, a data needs assessment study, background information on metadata, and a review of the benefits associated with the use of metadata. This complete package was planned and assembled with the

assumption that many organizations lacked both the knowledge and the proper tools for creating their own metadata files.

The Federal Geographic Data Committee (FGDC) contains 334 elements.

The Federal Geographic Data Committee (FGDC) approved the Content Standard for Digital Geospatial Metadata on June 8th, 1994. The metadata content standard consists of 334 different elements that can be used to provide information about a dataset. Though not all elements' fields need to be completed to provide adequate information about a dataset, it remains a dauntless task. Despite this, the FGDC standard provides a source for many organizations to adopt and modify to fit their needs. The FGDC offers metadata developers considerable latitude for their design considerations as they establish metadata guidelines to meet their goals while remaining committed to a standard that meets FGDC approval.

The Minnesota Governor's Council on Geographic Information Standards Committee[2] an entity formed in 1991 to provide leadership in the development, management, and use of geographic information in Minnesota, took the lead in establishing a metadata standard for Minnesota that complied with the FGDC standards. MetroGIS, the Minnesota Land Management Information Center (LMIC), and several other state agencies that included the Minnesota Department of Resources and the Pollution Control Agency, took the next steps by moving these metadata standards into the mainstream of GIS activities in the state and Twin Cities metro area.

The Minnesota metadata content standard contains only 101 elements.

The Minnesota Geographic Metadata Guidelines were approved as the standard in 1997, with the release of Version 1.1. These guidelines contain 101 elements, making metadata creation far more manageable, though time would tell if data producers would find this level acceptable. Considering that 30 to 50 fields require names, addresses, phone and fax numbers, emails, and some short answer entries, it would seem that this standard would persuade GIS data producers and users to create metadata for their datasets.

DataLogr is a metadata software tool that can be acquired through LMIC.

As part of the metadata implementation effort, a software program would need to be designed to conform to the Minnesota standard, and work effectively as a low cost tool for metadata creation. The software selected for the task was DataLogr@, a Windows-based metadata tool. IMAGIN, Inc., developed and designed the program to work effectively as a tool for the Minnesota metadata standard. IMAGIN represents a consortium of GIS users who organized in an effort to better develop and share data among government and educational institutions in Michigan. In 1998, an agreement between the Minnesota Land Management Information Center and IMAGIN allows LMIC to

---

[2] www.lmic.state.mn.us/gc/stds/metadata.html

distribute DataLogr to public organizations in Minnesota. The Metropolitan Council, the Forest Resources Council through the Minnesota DNR, and LMIC provided the financial support for the distribution of DataLogr.

Along with 101 elements or fields to enter information relevant to a dataset, DataLogr provides the functionality to operate on any Windows operating system, including 3.1, Windows 95/98, and Windows NT. The installation procedure involves only a couple of key strokes, and it opens a window that shows an edit screen that is used for entering and editing information entered into the fields.

DataLogr files can be used for Data Finder and the Minnesota Geographic Data Clearinghouse.

As part of support for DataLogr, web site development, and metadata file exchanges, LMIC developed a conversion tool that converts DataLogr files into several formats, including *HTML*, *SGML*, and *CSV*. The CSV format is required in order to display metadata and update existing metadata files in Data Finder and its web site. The Minnesota Geographic Data Clearinghouse[3], a node which provides a link between two database servers (Minnesota DNR and LMIC) and National Geospatial Data Clearinghouse, requires the HTML and SGML file formats. The National Geospatial Data Clearinghouse is associated with the National Spatial Data Infrastructure (NSDI). The NSDI is part of an FGDC-sponsored activity that supports a centralized clearinghouse. This site provides browse accessibility to participating organizations' spatial data and metadata across the world-wide web. A node provides the link to each organization's site.

Adding universal edit and edit window capabilities to DataLogr would greatly enhance the program's functionality.

The DataLogr program does come with some limitations that could hinder some efforts to implement metadata. The edit window is limited to one size and cannot be expanded to provide a better view of the information entered into the field. Its appearance is comparable to a DOS window. Another limitation is the ability to update multiple files. If a telephone number or contact name needed to be changed in 100 metadata files, each file would need to be opened separately in order to access the field that required the change.

Though there are these limitations, DataLogr does provide the necessary tool to start with metadata implementation at no cost to government and non-profit organizations, and the ability to import files will make certain that the metadata do not become stagnant due to the lack of updates.

As part of the effort to encourage data sharing among organizations in the metro area, a project was initiated in 1996 to assess information needs among participating members. Under the direction of the

---

[3] www.lmic.state.mn.us/chouse.html

MetroGIS Data Content Advisory Team, three sessions were held to survey and evaluate input received from 120 participants. Initially, as many as 870 needs were brought to the team's attention. Those numbers were reduced to 87 with the assistance of information management consultants. Finally, this number was reduced to top 13 information needs through the means of a survey that measured the importance of the information, and the dependence upon others for the data. High on the list of priorities were jurisdictional boundaries, parcels, land use, and other boundary information. In the context of metadata, the datasets that address each of the priority information needs would be important datasets to target for documentation, and be included in the strategy to implement metadata.

The top 13 data needs were used to target datasets for metadata compliance.

The last issue addressed, before beginning the metadata assault on organizations using GIS in the metro area, pertained to gathering background information on metadata and finding examples of benefits associated with metadata. These papers were combined into an information packet along with metadata brochures, a copy of DataLogr, its manual, and a software license agreement. This packet was distributed at meetings to GIS managers, technicians, planners, and administrators. DataLogr was not included in the metadata information packets during visits to private consulting firms because only public agencies and non-profit organizations could receive a free copy.

The metadata information packet contains materials on:

-Overview and benefits

-Standards

-Examples

-DataLogr software

-Contact Information

The following is the complete list of items enclosed in the metadata information packet that was distributed during meetings.

- An overview of metadata that Tanya Mayer, a GIS specialist at the Metropolitan Council, created for the project. The text font and writing style are very useful in removing the conventionality and banality associated with metadata. The document addresses most of the issues surrounding metadata and stresses the ease of metadata management and the many benefits. It also provides contact information and background information on the Data Finder project and MetroGIS;

- The NSDI's *The Value of Metadata* brochure that provides a federal perspective on metadata, though not as detailed as Tanya Mayer's brochure;

- The NSDI's *What are Metadata*? Sheet that provides a description of metadata in the context of the Nutrition Facts Food Label found on packaging;

- The *Minnesota Geographic Metadata Guidelines* (MGMG) document that provides the state's version of the FGDC Content

- Standards for Geospatial Metadata. Information includes the FGDC identifier, the FGDC element name, the Minnesota State Element Name, Field type, Domain, and Description for the 101 elements. The Description basically identifies the type of information needed to enter into the field;

- The *Detailed Descriptions of Metadata Elements* provides the same format as the MGMG document, though without the FGDC components. The document provides examples of entries that can be made in each field. The document is used for metadata workshops;

- Two hard copy examples of metadata files are provided in the information packet. These include metadata for the Municipal Urban Services Area (MUSA) dataset and the Washington County parcel dataset;

- Metadata and data sharing articles from magazines, journals, and various web sites were compiled to produce four case studies on the benefits of metadata in government, the private sector, the European Community, and for the data consumer. Examples of benefits were drawn from actual cases with the exception of the data consumer information sheet which provided examples of benefits from real and fictitious events;

- A copy of the DataLogr software on a 3.5" diskette was included with the packet for public and non-profit organizations. The diskette also contains HELP and EXPORT files, plus the 1990 Land-use metadata file to use as an example with DataLogr;

- A licensing agreement for the DataLogr software that must be returned to LMIC;

- The 40 page DataLogr manual;

- Business cards.

This compilation of metadata information and software would provide the key ingredients for making contact with metro area data producers and users, and giving them the capabilities to start producing their own metadata.

MARKETING METADATA

Contacts and Meetings

Counties, municipalities, regional organizations, and non-profit and profit private companies were visited.

The metadata marketing initiative cut across a wide range of public and private organizations, including county, city, regional, non-profit, and private enterprises. Because metro area counties produce and maintain the largest number of datasets, much of the attention was directed at their GIS staff. Several organizations in the private sector were also targeted because their companies provided GIS services to many small municipalities in the metro area.

The metadata information packets reached thirty departments.

In all, the metadata information packet reached more than 30 organizations and departments in the metro area, including all seven counties and several of their internal departments, eight cities, five private, three regional, two non-profit, and two university departments. Contact was made through email messages, telephone calls, and attendance at GIS user group meetings. The email and telephone were used to contact people to arrange for meetings at their offices. GIS user group meetings provided a good opportunity to disseminate metadata information to a crowd of 10 to 20 people at one time.

The visits were used to:

1)Distribute the information packets;

2)Discuss metadata and benefits;

3)Describe DataLogr and its use.

Private meetings were structured in a manner that allowed for considerable dialogue about metadata, inquiries into each organization's datasets, and the DataLogr software. Metadata issues covered in discussion included the definition of metadata and the Minnesota/FGDC standards, the benefits associated with using metadata, and the relationship between data sharing and metadata. The conversation then shifted to the types of datasets that each organization produced or planned for in the future. Finally, public and non-profit organizations received a free copy of the DataLogr software, while private organizations received the information packet, and contact names at LMIC for requesting the software. A review of the software provided the user with an overview that explained the ease of its installation and use, and DataLogr's export capabilities. The user was also made aware of the advantages of creating a template file that contained general information entered in the contact, metadata distributor, and metadata recorder fields. This information included names, addresses, phone and fax numbers and email addresses. An important point to get across to the user was that a template file contained standard information that could be saved and used from one metadata file to the next. Only information specific to the dataset would need to be added. The information entered into the elements, or fields, of the template file would account for more than one third of all the fields. Even among the remaining fields, many would require selection from a list, short answers, or no entries.

Metadata practices ranged from non-existent to active metadata documenting.

Meetings with organizations revealed a wide range of perceptive views towards metadata. A large number of people were aware of MetroGIS' activities with data sharing and its efforts to implement metadata. Among the 30 organizations that were contacted, a group of 12 were providing some form of documentation for their datasets. Within this group, four wrote short paragraph descriptions for their datasets. Four organizations provided background information that approached full compliance, but lacked some key elements, and were not storing the information in the proper format. The remaining organizations never provided an opportunity to see their metadata, so their descriptions of documentation procedures might place their efforts somewhere in the middle. It also seemed that many of these organizations had not documented all of their datasets. This group of 12 organizations can be best characterized as private sector (three), regional (one), county (four), and municipalities (three). At the time of writing this paper, Carver County became the fifth county to develop metadata, using DataLogr to create FGDC/Minnesota compliant metadata for their parcel dataset.

Discussions of metadata were linked to each organization's needs.

The larger share of the organizations knew little or nothing about metadata, including one public organization that was deciding whether to continue contracting out to private sources, or provide funding for an internal GIS operation.

Discussions were fashioned in a manner that would draw meeting participants to an understanding of the benefits associated with metadata and the ease of creating metadata files with DataLogr. An Ad lib approach was also utilized at each meeting so that the benefits of metadata and data sharing could be linked to the specific needs of each organization or department.

Findings:
1) Benefits were recognized;

2) No time or resources available;

3) Already providing non-compliant documentation that is meets their needs;

4) Metadata as another government mandate;

5) Boring;

In general, reaction to FGDC compliant metadata and DataLogr was mixed among the participants. Most recognized the value of metadata, or at least the potential; however, the implementation aspect was a stumbling block. Many felt that metadata implementation or maintenance would take away too much time from more important endeavors within their respective departments. As mentioned, some organizations had proceeded with their own documentation procedures, and believed that their methods and formats were adequate for their needs. Others hinted at the concern about mandated policies coming down from the federal government via the Metropolitan Council. Specifically, they believed that the existing standards could and probably would change again. These organizations preferred to wait for the dust to settle before venturing in the direction of compliance. There were also several organizations using a beta version of DataLogr, and some issues were raised regarding its functionality. The DataLogr

Organizations contemplating the use of a GIS are ready for metadata discussion because they can appreciate the cost and time effectiveness of metadata.

Organizations facing staff turnover and new staff are very receptive to metadata, especially if they were left with many undocumented datasets.

window cannot be expanded, so many of the fields are partially obscured for the viewer. DataLogr does not have the capability to run an algorithm to make universal changes in multiple files. If the area code number is changed, each metadata file must be accessed to edit the phone number fields.

Specific reactions to metadata depended in part on the amount of exposure or knowledge organizations possessed in the area of GIS and metadata. Most of the meeting participants were using a GIS, with the exception of one local community. It showed promise for future metadata compliance. The municipality considered the prospect of developing its own in-house GIS capabilities. Until recently, the municipality contracted out GIS projects to a private consulting firm. At the meeting, the metadata management and data sharing benefits were explained in detail to the participants. Consisting of an administrator and two planners, the meeting participants had no direct experience with a GIS; however, they expressed strong support for metadata because it was easy for them to appreciate the benefits that metadata had to offer, especially during staff turnover transitions, and locating datasets to augment and add value to in-house datasets.

The staff turnover issue provides an opportunity to segue to the another approach in marketing metadata to a receptive audience. On a regular basis, local newspapers and GIS job sites on the internet were monitored for new openings in the metro area. Especially important were postings for GIS managers. These types of openings were the metaphoric equivalent of "smoke on the horizon," which was based on the assumption that the preceding manager probably didn't take the steps to document datasets, thus creating a "fire emergency" for the incumbent. The reactions to metadata at the meetings reinforced this belief, and this was confirmed during transitions at two metro area organizations. In both situations, the incumbents were left with no information about existing datasets at the site. The time and effort involved in locating and identifying these datasets was considerable, and provided a strong incentive for the new GIS managers to recognize the importance of documenting datasets. As expected, there was a resounding acceptance of metadata conveyed when contact was made with the two managers. Both attended the metadata workshop, and implementation has started at one site, and the other has promised to start soon.

Reactions among other organizations that had not delved into metadata, or dataset documentation, could be described as ambivalent and often accompanied with some laughter. Usually the first questions posed to all meeting participants regarded the methods employed to document information about their datasets. Surprisingly, at times with more than

Some use only their memories to retain dataset information and had no interest in metadata.

one in attendance, a spontaneous laugh would break out in unison, followed with the comment "in our heads." A metadata company called "Enabling Technology," uses the logo "Storing Metadata in Your Heads?" and these reactions provided a convincing testimony to the Enabling Technology's perception of the metadata world. Though these individuals and others appeared receptive to the benefits associated with metadata, time and resources were mentioned as serious concerns that would exclude them from addressing metadata any time soon. There were several instances when staff members from a couple of organizations doubted the value of metadata. These were organizations with only one or two people on staff who dealt with all the GIS responsibilities.

Regional agencies and small in-house operations were receptive to using metadata because of the potential for acquiring and sharing information about datasets with metro area jurisdictions.

The remaining GIS users, also unfamiliar with metadata, appeared to be very receptive to metadata, though their organizations had only a few or no datasets to document. These organizations are best characterized as regional entities serving the metro area and two departments at the University of Minnesota. None of these entities would be considered a prolific data producer. They are data users, so data sharing and metadata are important for their organizations. Metadata can serve to provide the means for them to assess the availability and quality of datasets from other organizations, and to document their own datasets for internal management and provide dataset information and accessibility to users throughout the region.

Among the organizations involved with dataset documentation, the private sector takes the lead.

To date, organizations that are implementing metadata or dataset documentation can be characterized as representing three private GIS consulting firms, one regional and one non-profit organization, four metro area counties, and three municipalities. Conversations with GIS managers at both the private profit and non-profit organizations revealed that they have been creating metadata, or writing detailed descriptions for their datasets. Several mentioned that this policy had been in place for at least several years, or since the opening of their company. The shared opinion among the managers, except for the non-profit was the cost of the DataLogr software. It seemed to be a question of principle based on the fact that public and non-profit were not charged for the license. The private firms also provided service to numerous clients in the public sector, primarily municipalities; therefore, the case was also made that a private firm shouldn't be charged for a software tool that their public client can acquire gratis. On the positive side, these private organizations provide GIS services to local government agencies, and some form of data documentation is included in the service. These private consultants could also serve as a useful resource to develop metadata for the municipalities that operate small GIS departments and produce their own datasets. A contractual agreement could be reached that permits the consultant to assist in the

Private consulting firms provide a valuable service to small agencies that lack the resources to produce their own metadata.

development of metadata as an additional, or value-added service. Municipalities with in-house GIS operations, but limited resources or a small staff, may welcome this opportunity to out-source or receive training with metadata development so that it frees their staff to focus on GIS applications.

Some organizations made the case for starting with a data catalogue rather than fully compliant metadata.

The one non-profit organization that agreed to meet had also worked with metadata for several years, but argued in favor of a data catalogue, or a condensed version of metadata. The case was made that a data catalogue was adequate for tracking and managing in-house datasets. If there were a need to exchange data, a data catalogue could be modified to a format that produces a FGDC/Minnesota compliant metadata file. This issue has merit, and should be considered as part of the implementation strategy. One problem is that it falls short of the data sharing strategy that would require sufficient information for potential users to effectively evaluate a dataset. The user would need to contact the data developer to inquire about the missing information, and part of the metadata benefit is lost because time must be dedicated to taking phone calls to answer questions about a dataset. In the end, a data catalogue is an effective approach to internal management, but not necessarily for promoting data sharing.

Organizations that have started their own documentation procedures are reluctant to adopt a new metadata standard from an external source.

Among the four counties visited, one opined the same philosophy about data documentation as the non-profit organization. The county took a lukewarm stance to the FGDC/Minnesota standard, and didn't appear interested in changing its policy. Their approach was to write a one or two paragraph description as their version of metadata for each dataset in an HyperTextMarkup Language (HTML) format, and distribute it with requested datasets. On a periodic basis the county also distributed their version of metadata to potential users so as to inform them about the development of new datasets. The county did make an effort to create metadata files using a beta version of DataLogr, and as many as 19 datasets were documented. An intern had been assigned the task of creating the metadata files; however, a review of the metadata revealed that more information was required to make their metadata compliant.

Several organizations have made serious attempts to create compliant metadata, but lack the resources to bring their efforts to fruition.

The other three counties took important steps to produce their own FGDC/Minnesota compliant metadata, though a quick review of their software formats and metadata contents revealed that some changes would be required to attain full compliance. One county adopted the FGDC/Minnesota standard format and created metadata using MS Access. This county was especially interested in selling its datasets, so metadata could prove to be a useful data catalogue for marketing. Their metadata records for the respective datasets would need to be extracted from the database file and separated into individual DataLogr files. This would provide compliant metadata that could be displayed on the

Data Finder web site. The reason stated for using MS Access was that a metadata tool wasn't available at the time when their department wanted to start documenting information on their datasets, and MS Access provided a tool for running queries to locate specific datasets.

Some organizations have mixed metadata records into one file, making extraction difficult.

A meeting with the second county in this group revealed that they had used a beta version of DataLogr to create metadata. The dataset information contained in the metadata file was detailed and quite complete, though one metadata file contained information for all of their datasets, which were associated with their parcel dataset. This presented the same problem as encountered with the county using the MS Access program. Again, documentation for each dataset would need to be extracted and copied to separate DataLogr files.

The third county would partake in a project that would involve a contract with the Metropolitan Council to use one of its metadata assistant to assume the responsibility of reviewing the county's existing metadata and take steps to create metadata for the other county GIS datasets. The department responsible for GIS data development had already created detailed metadata files that were FGDC/Minnesota compliant, though the metadata file for their parcel dataset had information about other datasets embedded in the file. These datasets represented AutoCAD layers, which were part of a data dictionary file that contained a description for each layer.

Three willing participants at the municipality level, with two making serious attempts.

The three municipalities can be characterized as willing participants in metadata development. One municipality had created only one metadata file for a zoning dataset. Another used MS Word to store dataset information, and examining the file revealed very good documentation. The third municipality used an Apple spreadsheet program that included a brief, general description for each dataset.

One regional agency expressed great reluctance to go with fully compliant metadata.

The contact at the regional organization mentioned that dataset information was documented at their department; however, no interest was expressed in using DataLogr or creating FGDC/Minnesota compliant metadata, so it's assumed that their approach wasn't particularly robust.

A discussion of metadata and a demonstration at a user's group meeting did not generate interest.

One occasion provided the opportunity to meet with representatives from about half a dozen local governments at one site. The people in attendance were members of a GIS users group, and one its members had extended an invitation to the Metropolitan Council to send its metadata assistant to address the issue of metadata before the group. Metadata information packets were placed at the table for people to collect, and time was allotted at the end of the meeting for a discussion about metadata and a demonstration of DataLogr. The DataLogr

demonstration proved effective at meeting with county representatives at one site; however, with the meeting convening towards the end of the day, several people departed before the metadata discussion, and several in the group teetered between sleep and metadata. Only one person had taken an information packet.

An Assessment of Responses and Reactions to Metadata

A majority among the 30 organizations contacted during the first stage of this effort expressed an unwillingness to shoulder the responsibility of creating fully compliant metadata.

An overall assessment of the feedback received from the meetings, telephone calls, and emails revealed two dominant views among professionals in the Twin Cities metro area. A majority of professionals in the field appeared unwilling to shoulder the responsibility of producing FGDC/Minnesota compliant metadata. Another group lacked the resources to effectively address the issue of metadata compliance. Most recognized the benefits associated with using metadata to document information about their datasets. Strong support was voiced for using metadata to manage internal datasets. Data sharing benefits were viewed with suspicion among some participants and indifference among others. A minority found little need for metadata, especially metadata conforming to the FGDC/Minnesota standards.

Work priorities dominated concerns about metadata, especially at the departmental level for larger organizations.

The rationale behind many of the indifferent views toward metadata reflected on an organization's operations, priorities, and perceptions. This was especially apparent at the department level of most organizations, especially among surveyors and engineers. Some would devise their own system for documenting their datasets, and the method proved adequate for their purposes, so there was no need to go beyond that scope. Typically these departments maintained greater control of their datasets. Their staff was adept at using a GIS for developing datasets, and the staff most often assumed the responsibilities of running GIS applications and producing maps. Their perspective was that other departments lacked staff members with proficient GIS skills, so there would be little or no internal demand for any datasets, thus making metadata less relevant. If metadata were developed, the format was structured to fit departmental needs. Internal or external requests for data would be expected to accept their version of metadata.

The perception was that no internal departments would share datasets, let alone metadata.

Close-knit departments and one person operations trust in themselves that the dataset information is safe in their heads.

The few departments that lacked any interest in metadata represented one-person operations, or were composed of close knit groups that had worked together for some time. The one-person department usually answers to no one. The group that relies on trust and camaraderie believes that no one amongst them would consider leaving. As long as no one makes a career change, abruptly departs for another job, or takes

a long vacation to Antarctica, the information is always there. It's just a matter of asking the person most familiar with the dataset a few questions.

Some organizations are concerned about revealing internal datasets through data sharing, data scrutiny, and data liability issues.

Another possible reason for some organizations to avoid metadata is the concern over public scrutiny of their datasets. Producing metadata takes an organization one step closer to sharing information about their datasets. Though there are metadata elements that list liability clauses and address access and user constraints, individuals within an organization may not feel comfortable with releasing dataset information to the public domain. During several meetings, the spatial accuracy concern was brought to the surface. Despite reassurances that metadata could be used to describe these limitations, this issue appeared to remain a concern.

Despite the ambivalence about metadata, a majority of the organizations appeared to favor the development of compliant metadata, if provided the means.

Despite these attitudes, a majority of organizations and individuals expressed support for data sharing and producing FGDC/Minnesota compliant metadata; however, most of them lack the resources to effectively initiate or update their metadata. As a means to entice organizations towards implementing metadata, the metadata template method was offered as a solution to their time and costs concerns. This approach didn't convert any of the doubters, especially those people aware that there were still a number of elements in the standard that required considerably more than short answer/option entries. Considerable time would still need to be dedicated to providing spatial reference information, identifying and defining attributes, and providing data sources and information on the processing steps taken to produce the dataset.

Among the organizations with metadata or documentation, the case was made that a cut and paste approach could produce compliant metadata. The ease of copying text from MS Word, MS Access, or the beta version to DataLogr to DataLogr didn't muster much cooperation either.

PROJECT EVALUATION AND IMPLEMENTATION OF NEW STRATEGIES

Evaluation of First Project Stage

For many organizations, data development was still their GIS priority, so metadata were not entered.

The contacts made with organizations, departments, and individuals revealed that the FGDC/Minnesota metadata standards were not considered a high priority. The reasons ranged from resources to perceptions, priorities, and dealing with the least attractive aspect of GIS and data development. A number of organizations were also in the beginning stages of data development, so it is difficult to ascertain if metadata compliance will become a priority in the future. A review of

Most articles in journals and magazines focus on the conceptual rather than implementation aspects of metadata.

relevant journals, proceedings, and magazine articles revealed that MetroGIS was at the vanguard of this metadata implementation endeavor. Most of the pieces on metadata revolved around the conceptual side rather than implementation. This probably explains some of the problems encountered during the first stage of this project. Many articles suggested that metadata was still going through an evolutionary process, with a prevailing attitude that data producers and users will recognize the benefits of metadata, and begin implementation without any nudging or incentives. It makes sense that the benefits of metadata would provide a strong incentive for implementing such a procedure; however, it doesn't take into account the needs of human nature.

One article, relevant to metadata implementation, examined needs that motivated human behavior.

The literature search opened a door to a paper that was presented a 1996 IEEE conference in Maryland (Callahan, Johnson, and Shelley, 1996). The paper, *Dataset Publishing- A Means to Motivate Metadata Entry*, addressed the issue of metadata implementation and the factors that influenced an individual's behavior. The three authors proceeded to cite a behavioral study (Murray, 1938) that identified primary and secondary factors that influenced or motivated an individual's behavior. The primary factors that affected behavior were linked to biological requirements such as the need for air, water, food, and sex. The secondary factors, or needs, were not linked to the physical or biological needs of an individual, and pertained to learned or acquired psychological factors. The authors went on to list several of the 28 needs that Murray had identified in his study, and indicate that these needs are relevant to metadata implementation. These include the need for acquisition (e.g., monetary reward), achievement, recognition, exhibition, autonomy, affiliation, order, cognisance, and exposition.

The key task is translating needs into motivational tools that influence an individual's choice between metadata implementation and other competing task.

Another portion of their paper discusses the realm of the NPI Theory that describes the factors that influence an individual's behavior within an organization. Specifically, the theory deals with, as stated in the text, "why the individual chooses certain alternative courses of action in preference to others, and thus it might properly be called a theory of choice behavior." In order to achieve some success with implementing metadata, it is necessary to identify metadata's benefits to an organization so that these can be translated into motivational tools that influence an individual's decision process in determining the amount of time and energy to expend towards metadata production against other competing tasks.

The IEEE article reinforced some personal perceptions made during the "introduction to metadata" meetings with the metro area GIS organizations. Many of the meeting's participants made it apparent that metadata didn't meet any of the four major needs criteria, or for that

Order, autonomy, affiliation, recognition, acquisition, and achievement are recognized as needs that could be applied to metadata implementation strategies for metro area organizations.

matter the second group of 28 needs. During the course of the meetings, several of the needs were mentioned to the participants, though for many, a strong case couldn't be made in favor of metadata. Metadata offer an organization a considerable amount of order, autonomy, affiliation, recognition, acquisition, and sense of achievement.

Providing order seemed to offer the strongest case for using metadata because an organization could manage its datasets effectively. Dataset maintenance would assure an organization's control despite staff turnover. Metadata are relevant to an organization's autonomy, affiliation, and acquisition. Some degree of autonomy is possible for an organization if it uses metadata because a potential user will not have direct access to the organization's datasets. A metadata file also provides information for a potential user to evaluate datasets; therefore, eliminating or reducing the amount of time that individuals in an organization must spend on the telephone with potential users. A metadata file can also divert potential users to the contact or a data download site, if a data transfer is necessary. These latter two conditions provide the sort of autonomy that frees individuals in a department from administrative duties, and provides the time to focus on more interesting and enterprising tasks involving GIS. This is under the assumption that there is a preference of creating metadata over conducting administrative tasks.

Affiliation and acquisition are self-apparent in that metadata create the opportunity to acquire and exchange data between organizations. The datasets can add value to an organization's existing datasets to apply towards new applications and reduce the costs of capturing and processing the data.

Affiliation and recognition can be considered as devices that can be used to promote peer pressure as a means to encourage more individual and organization participation to engage in metadata development.

These recognized human needs merit serious evaluation that can be applied to short and long-term strategies. Human actions fell short of expected goals, thus an approach must be found that translates the message of metadata benefits into human action. For various reasons, many organizations rejected the metadata benefits concept in favor of other competing tasks, or because the need for documentation had been met using their own version or method of data documentation. The MetroGIS' strategies to be implemented would need to take into account the diverse cultures in the GIS world, metadata and organizations' and MetroGIS' perceptions of metadata, and the problems or limitations that discouraged each organization from

implementing metadata. In effect, MetroGIS would need a new metadata recipe that would please the palates among the metro area GIS connoisseurs and dilettantes.

Immediate Approaches to Metadata Implementation Strategies

Recommendation:
Target organizations with large number of datasets, especially those with legacy datasets.

The metadata implementation process must encompass both short and long-term strategies in order to be effective. The short-term approach to the problem must accomplish the tasks of providing support to organizations interested in implementing metadata, and both pressure and assistance to those that have created metadata, that does not adhere to the FGDC/Minnesota standard. Especially important in this endeavor is to target organizations that have a large number of legacy datasets. As time passes, these datasets could become obsolete without proper documentation.

Recommendation:
Meet administrators and GIS managers during transition phases to discuss metadata and incorporating implementation into job descriptions and responsibilities.

As mentioned briefly, visiting organizations that plan to start in-house GIS operations provides an opportunity to meet with administrators and discuss the benefits of metadata, and the need to include implementation and maintenance as part of job descriptions and reviews for their GIS managers and staff. The second is to watch for transitions through job announcements for GIS managers or specialists. Most likely these incumbents will find that their predecessors avoided dataset documentation.

Another important strategy that can be employed with metadata implementation is to offer metadata workshops. MetroGIS and LMIC have sponsored half-day metadata workshops at the annual GIS/LIS Minnesota conference. The workshop provides participants the opportunity to learn about metadata and the software tool, DataLogr. Participants also spend some time working with DataLogr, and are encouraged to bring dataset documentation to enter into the program; however, very few people take advantage of the opportunity for a variety of reasons.

Recommendation:
Organize half-day metadata workshops, with time for participants to produce their own template files and enter information for their existing datasets.

A similar approach was taken in January, 1999 when twelve organizations were invited to send representatives to a Metropolitan Council sponsored, hands-on three-hour metadata training session. The workshop followed many of the same procedures as at the GIS/LIS conference; however, there was less emphasis on the theoretical side of metadata, and more on creating metadata, especially the template file. Participants were also strongly encouraged to bring documentation for one of their important datasets. A list and description sheet was sent to participants prior to the conference. The sheet provided suggestions about which dataset information should be collected and brought to the workshop. MS Access and MS Word were also loaded onto the

computers for those participants that had dataset documentation, but not compliant to the FGDC/Minnesota standards.

Attendance was very good for the workshop. There was only one cancellation, and those in attendance worked on the metadata template version. Almost everyone brought documentation to the workshop as well, and some time was dedicated to entering information into DataLogr. The workshop received good reviews on the questionnaire that was distributed. An overall assessment would have probably called for more time to enter information into DataLogr because all of the participants had received the metadata information packet during previous meetings, thus requiring less background information about metadata.

Recommendation:
Offer additional metadata workshops on an annual or semi-annual basis.

The success of the workshop revealed the need to offer subsequent workshop on a periodic basis. The workshop sets off a block of time for participants to escape the office and focus on metadata. The workshop also offers the participants the opportunity to discover that metadata implementation is not a difficult burden.

Recommendation:
Provide interns to organizations interested in updating or creating metadata files.

The one element that didn't succeed with the workshop was an attempt to provide student interns to assist organizations that lacked the resources, but strove to develop metadata. Several months prior to the workshop, MetroGIS approached a local private college about the prospect of inviting a couple of student interns to assist with the metadata project. As a requirement for meeting their graduation commitments, students in the geography department would be required to dedicate about 140 hours to internship activities during the course of a semester. Initially, two students had expressed interest in participating; however, they encountered some serious time conflicts in their spring schedule, and couldn't participate in the project.

Recommendation:
Require organizations, wishing to work with interns, to send professionals and interns to the workshop.

Despite this minor setback, a door was open to cooperation between MetroGIS and a local college to offer opportunities to student interns and assistance to data producing organizations in the metro area. Future plans will be to encourage students to get involved as interns, recognizing that their efforts will provide access to prospective employers, plus experience with metadata. The students would become familiar with metadata, and receive training in the use of DataLogr and the use of some GIS and metadata media software packages.

The interns would be present at the metadata workshops, to provide assistance and meet with the participants. The interns' services would be available only to those attending the workshops because both sides will have been exposed to metadata and understand the same

terminology that will facilitate information exchanges between the interns and the organizations.

After the workshop, MetroGIS would target organizations holding the most legacy datasets, and then strongly encourage them to work with the interns. Once selected, the organizations and interns could coordinate efforts to work together. Interns could interview key staff members to acquire information about a dataset. Another option would be for an organization to give the interns the datasets that require documentation, and using a GIS, the interns could extract information for the metadata, or ask more specific questions that require less time to answer. If an organization has documented datasets, the document files could be handed over to the interns, so that the information can be copied from the organization's medium to individual DataLogr files.

As a short-term goal, this strategy provides an opportunity for organizations to gain some important ground in their efforts to document existing datasets. The metadata workshop and interaction with the interns also offer the chance for staff members to become more familiar with the metadata elements, and the information required to produce a FGDC/Minnesota compliant metadata file.

Though the opportunity never materialized for the interns to work with the organizations, the experience that could be used to implement such an endeavor was acquired when MetroGIS collaborated with a metro area county to produce metadata for their datasets.

A Short-term Metadata Implementation Strategy: Dealing with a Backlog

MetroGIS provided assistance to Washington County as part of an effort to update its metadata.

As part of the metadata implementation project, MetroGIS and a metro area county, Washington County, collaborated to share resources in an effort to update their dataset documentation. The Surveyor's department, their leading geo-spatial data producer, had already made strides in producing quality metadata, especially for their parcel dataset. The objective would be to create metadata for other datasets that had been produced within the county, as well as for other county datasets that other agencies had produced over the course of the past several years.

The first meeting with Washington County officials revealed a transition in progress. It could very well represent a trend within the industry, and make metadata even more important. As mentioned, traditionally the Surveyor's department was responsible for developing datasets. This included all GIS operations as well. The transition would create a new IS or Information Services department that would

work more with GIS applications using their own datasets, those from the Surveyor's department, and datasets from other agencies.

The existence of two GIS departments in one organization, creates an internal need for metadata.

This transition provided even more incentive for the county to develop metadata. It would also provide an opportunity for MetroGIS to draw upon the experience of working with several departments in an effort to create metadata.

A total of 39 datasets and subsets were targeted for metadata.

The first meetings were held with representatives from the county planning and IS departments, and the first order of business was to identify and target datasets for metadata documentation. These included datasets embedded in the parcel dataset, which contained AutoCAD layers that hadn't been extracted, but could represent separate datasets. Initially, a total of 39 datasets and subsets were targeted for metadata. The subsets represented separate files that were linked to projects, or components of map atlases. As many as 24 of the datasets had metadata files; however, most contained inadequate, or missing information.

A template file comprised of names, addresses, phone numbers, access and user constraint information and liability clauses, was created.

The first ones targeted were those datasets without metadata files. First a template file was completed, which contained information about the contact, distributor, and person responsible for the metadata file. For simplicity, one person from the IS Department was selected to represent all three metadata elements. Address, phone and fax numbers, plus email address were collected via the email. Access and user constraint information and liability clauses were extracted from the parcel metadata file along with the ordering instructions, spatial reference information and bounding coordinates. The parcel metadata file provided complete information about the coordinate system, and all datasets originating from the county or other agencies were referenced to the county's coordinate system.

Dataset information was acquired through a series of questions using email messages, and a review of datasets sent to the Metropolitan Council.

Dataset information was also acquired through a series of steps involving email messages and reviewing datasets that were made available to the Metropolitan Council. The IS department was offered two options. They could receive a wide range of questions about the datasets through phone calls, interviews, and emails, or the department could send the datasets to the Metropolitan Council for review, which would follow with more short, concise questions. The IS department preferred the latter, and when possible, sent the datasets as attached files to email messages. The files consisted of compressed Arc/Info export and shapefiles. After the files were received, they were uncompressed and then opened and viewed in ArcView.

Access to the datasets provided an effective means for extracting information about a dataset, especially its attributes. The attributes

were usually represented as codes that could be located in the table or map legend. The codes could be written down in an email message with a request asking for code definitions and attribute descriptions. Dates, abstract, purpose information, dataset environment, vendor object types, file sizes, and information about other elements could also be extracted from the datasets, though some inferences were necessary for the abstract and purpose statements.

The most challenging task was acquiring information about the data source and the processing steps to create the dataset.

The remaining important elements required more time and questions to acquire the all the necessary information. Most important was information about the data sources and the processing steps (lineage field) used to produce the dataset. Usually, the county or agency provided descriptions that were written at the time the dataset was produced, or completed. If an agency had produced the dataset, a description of the sources and processes was written on a letter that accompanied the dataset to the county. At times the information was sufficient for writing to the metadata file, and other times it required further investigation with follow-up questions to the dataset's originator.

The datasets with metadata were reviewed for quality and content.

The same procedures were taken for datasets with metadata files, though the files were first reviewed for quality and content. In some cases, existing information from other metadata files could be used to replace inaccurate information found in a metadata file. As an example, the bounding coordinates in 24 metadata files were listed in UTM coordinates rather than latitude/longitude. The correct coordinates were extracted from the parcel metadata file, and used to replace the incorrect entries. Without this reference source, it would have been possible to determine the coordinates using a GIS, and pulling in one of the county maps to get the coordinates off the screen with the use of a mouse.

As with undocumented datasets, questions were posed to the Washington County contacts, or accurate information was taken from existing metadata.

Again, as with the undocumented datasets, specific questions were posed to the contact in the IS department and the agencies. In some cases, there was no additional information to provide, or the information was broad and general, especially when associated with the lineage field.

When all the metadata files were completed, they were returned to the County for a final review.

When these metadata files were completed, they were returned to the IS department for a final review to inspect for inaccurate entries missing or inadequate information. Inconsistencies encountered in the metadata were brought to the attention of MetroGIS and resolved at a meeting. A metadata tracking/status sheet in an MS Excel format was created for departmental use as well. The sheet contains the names of the datasets

and their respective metadata files, status columns to provide updates on the status of metadata files, a dataset availability column that indicates if datasets are available for documentation, and a column that lists the internal contact names. The contact names represent sources to reach if questions should arise with regards to metadata or a dataset.

A review of the metadata for the parcel dataset and its data dictionary revealed 12 data subsets that could be documented.

Upon completion of the metadata for the targeted datasets, attention turned to the parcel dataset. A review of the AutoCAD data dictionary revealed that there were 12 subsets that could be extracted and used as GIS or CAD datasets. This list was brought to the attention of the IS and the Surveyor's departments. The IS department regarded this list as a priority for metadata; however, the surveyors didn't concur because no requests had been made to extract the subsets to use for other applications. They didn't see a need to expend time and resources on these subsets, unless requested, and for this reason documentation wasn't necessary at the time. The priorities between the data developers and data users needed to be resolved before proceeding any further.

A meeting between departments established that the documentation of the 12 data subsets would be dealt with at a later date.

Consequently, a meeting was called and the issues were presented and subsequently resolved. The meeting revealed that one of the CAD layers had been extracted from the parcel data dictionary, and used for a GIS application. The metadata file for the parcel dataset would continue to serve as back-up documentation for the other subsets, and it's assumed that the metadata files would be produced for each of these subsets at the time when extraction is deemed necessary. The surveyors also mentioned the availability of three other metadata files, including one that had been initially targeted for metadata. The three files were in an HTML file format, and were sent attached to an email several days after the meeting. The files were to be reviewed, then copied from the HTML files to individual DataLogr files for final editing. The three metadata files were well done and required only a few additions and changes. The tasks were completed and the files were returned to the surveyors for a final review.

Finding:
The 145 hours required to complete 43 compliant metadata files are broken down as follows:

1) 10 hours for meetings;
2) 20 hours for metadata reviews;
3) 60 hours for viewing datasets and creating new metadata files;

A total of 145 hours were expended towards creating 43 FGDC/Minnesota compliant metadata files for the county, of which 15 were new files. Specifically, 10 hours were used for meetings, 20 hours for metadata and the parcel metadata review, another 60 hours for viewing datasets in ArcView and creating new metadata files, 30 hours for editing existing metadata files, 15 hours for exchanging questions and answers via emails and phone calls, and another 10 hours for final metadata file reviews. Broken down, about 3.37 hours were dedicated to the development of each metadata file. Excluding meetings, the number time drops to about 3.1 hours. The number of hours spent entering information into DataLogr to create a new metadata file averaged about 1.4 hours creating a new metadata file in DataLogr.

4) 30 hours for editing existing metadata;
5) 15 hours for exchanging information; and
6) 10 hours for the final metadata reviews.
On average, it required about 3.1 hours to produce a complete metadata file. Less than half this time was dedicated to entering all information into a new metadata file.

The experience and knowledge acquired from the Washington County metadata project will prove useful for organizations with a backlog of undocumented datasets.

This is comparable to a figure cited in a study (Lynne Bly & Associates, 1997) that provided an estimate of 1.5 hours for creating each file. The 1.4 hour estimate for the Washington County project does not include the amount of time dedicated to acquiring dataset information and reviewing the completed metadata files. Development of new metadata files accounted for about 40 to 45 percent of project's time, while editing existing files amounted to about 20 to 25 percent. Both figure ranges take into account the time committed to acquiring information about the datasets. An additional 15 percent of the time was dedicated to initial and post-processing reviews of new and existing metadata files.

The experience acquired from the Washington County metadata project produced many constructive procedures that could be applied to future cooperative efforts between the Metropolitan Council and participating organizations. The first and foremost issue that must be addressed is communication and assignment of responsibilities. Many organizations face the same dynamics as encountered with this project. More than one department may be involved with a GIS, with one department responsible as the data developer, while another plays the role as both a dataset user and developer. In many cases, surveying and engineering departments assume the task of developing datasets. Their skills provide spatially accurate datasets. Other departments, such as planning, economic development, and environmental, can be categorized as users and developers. Their staff will develop applications for the datasets and create new datasets in the process. Bringing these two interests together along with their administrators provides the opportunity to address all the issues so that staff members or interns know whom to contact for questions about specific datasets. Pulling all sides together provides a reminder about the importance of metadata to each of the groups. The dataset developers should see metadata as a management tool, while metadata serve as both a management tool and data access mechanism for the planners. Administrators should view metadata as the tool that can bridge communication gaps between departments, and provide revenue and resource security against the event of staff turnover and transitions. Getting all these parties to cooperate with metadata development will facilitate the metadata documentation process for the intern. As described previously, the tracking, or status sheet approach will prove useful as a management tool during the process.

Recommendation:
The intern should acquire as much dataset information as possible without causing too much disruption.

The intern's role with any cooperative endeavor is to acquire as much information as possible without too much disruption to a department's daily activities. The county metadata project revealed that an effective approach was to gain access to the datasets. As described previously, this provides the opportunity for the intern to extract information

directly from the dataset without requiring assistance from the contact. Questions to the contact can be made more specific too, and this method usually receives a faster response. Reducing the time that the dataset contact must dedicate to answering questions also saves the department additional cost and resources.

Recommendation:
A participating organization must make a final review of the metadata.

Finally, the intern must be responsible for assuring that the organization's contacts have reviewed the completed metadata files for quality and content. The files must also be reviewed to guarantee that the FGDC/Minnesota standard guidelines are met. The participating organization will also be encouraged to provide the metadata files to the Data Finder web site.

Recommendation:
Hiring interns, combined with workshops, offer the best solution to deal with the backlogs of existing undocumented datasets.

This hands-on approach appears to be the only effective short-term solution to metadata implementation for the metro area. Combined with meetings, information packets, and other marketing schemes, this intern resource approach, combined with the metadata workshops, seems to provide the best solution, especially among organizations with large numbers of datasets and few resources. In situations when organizations hire their own interns, MetroGIS must offer its experiences and resources to training and assisting the interns if their duties include metadata development. Metro area organizations may also wish to pool their resources to work with MetroGIS and the Metropolitan Council to provide interns for future metadata implementation projects.

Long-Term Strategies for Metadata Implementation

Recommendation:
Academia must introduce metadata to their students.

Any long-term approach to metadata implementation requires cooperation with educational institutions. Federal and state agencies involved with data access issues and metadata must establish contact with the respective departments at universities, colleges, and technical schools to discuss measures to incorporate metadata into classroom material. Based on the diverse group of professionals working in the field, a multi-disciplinary approach must be considered in this strategy. This range includes the fields of geography, urban planning, civil engineering and surveying, business marketing, or any discipline that has the potential for handling geospatial information. In many cases, exposure to metadata can be limited to one or two lectures and a class or lab assignment that involves the documentation of a dataset.

Metadata implementation strategies at the local scene present many challenges that require an array of tools. Recalling the section about needs that affect human behavioral patterns, there are several needs that can be considered when designing an implementation strategy.

Metadata compliance can be tied directly to an employee's work duties, using job descriptions and performance reviews. This approach goes directly to the individual, and can represent a GIS manager, another staff person, or all the people in a department. The individual(s) must be made accountable for producing metadata with the understanding that there is organizational support for this endeavor. Their efforts can be acknowledged through recognition, favorable pay increases, and other perks.

Unfortunately, including metadata responsibilities into a job description has not become an ubiquitous concept among the many managers in North America. A recent search that was run on the GIS Job Clearinghouse site revealed that only two jobs out of almost 200 postings had included metadata in the ad. For this reason, it appears that this approach still represents a long-term solution.

If these approaches are to be expected to succeed, then an organization's administrator may need to be included in the metadata implementation effort. The visit with the municipality that was considering in-house GIS provided the only opportunity to meet with an administrator. All other contacts were with GIS data producers and users and their managers. These people had control over their own destinies and their departments. The pressure to implement or produce FGDC/Minnesota compliant metadata fell into their hands, and most measured the risks of eschewing metadata against other competing demands in their departments. As mentioned, metadata haven't faired very well in the contest.

Though administrators may not understand all the concepts associated with datasets, GIS, and metadata, they can easily understand the importance of metadata in the context of cost and management. Administrators are keenly aware of the effects associated with staff turnover, and explaining metadata in the context of maintaining continuity through transition provides an enticing solution. A case can also be made for metadata when applying it to data sharing as a source for access to other datasets that could augment and add value to in-house datasets. An administrator should be convinced that both outgoing and incoming datasets have metadata files accompanying them. This approach can be implemented through a contract that has a metadata clause written into the agreement.

GIS managers represent important contacts as well, though it's difficult to know when to include an administrator in the process. Recommended as well for a short-term strategy, the best opportunity to contact managers is when a transition is taking place. The experience

a receptive audience for metadata completion.

of expending time to locate and identify undocumented datasets provides a convincing testimony for an incumbent manager or technician during a transition. The experience can easily provide a convert to metadata.

Recommendation:
Target prominent organizations to develop and display their metadata on Data Finder to encourage other organizations to comply.

Another motivational tool is to apply peer pressure to organizations, especially those that use their own documentation methods that do not comply with the FGDC/Minnesota standards. If the short-term student intern strategy is effectively implemented among organizations with large numbers of datasets, this will provide a valuable resource that can be used to encourage others to participate. Should the Data Finder web site carry hundreds of metadata files from participating organizations, the non-compliant ones may desire to become affiliated with this movement, especially if data sharing becomes commonplace. Affiliation is another need that was identified as an influence on human behavior.

Recommendation:
Collaborate with private consulting firms to assure that they create compliant metadata for their clients.

As mentioned briefly in a previous comment, another tangible and important resource for metadata development is the private sector. Most of the private sector has recognized the importance of metadata, and has taken steps to produce detailed documentation for managing their datasets, and those developed for clients. In the metro area, a significant number of municipalities use private consulting firms to develop their datasets. This relationship provides a valuable opportunity for the Metropolitan Council and other agencies to work with these private firms to assure that clauses are being written into contract agreements, guaranteeing that private firms will produce compliant metadata for the municipalities.

Recommendation:
Remind MetroGIS participants of the importance of metadata for data development and data sharing.

Among metro area organizations in the public sector, metadata implementation represents an intricate part of the MetroGIS data and cost sharing agreements. Organizations volunteer their efforts to comply with provisions in the agreements to fund data development and data sharing among producers and users. Metadata provide the impetus to offer the means for these groups to interact so that one motivates the other in an effort to encourage and sustain data development and use. As more datasets are produced, their values increase for developers and users because more of the geo-spatial elements of the metro area are converted to a digital medium, and the users can develop more applications to address the issues and concerns of regional development and day-to-day planning and management. Metadata provide the link that can turn this possibility into reality.

The audience of data producers will also offer solutions for the future. Some have suggested that the metadata community consider using data catalogues as a medium for storing dataset information. If a request is

Recommendation:
Review and update metadata
and the software tools,
especially DataLogr.

made for the dataset, the catalogue could be updated so that the file meets the metadata guidelines. This approach falls short of the objectives; however, it may require this step-by-step process to get more organizations involved. Conversations with people have revealed their angst over changes in the FGDC/Minnesota guidelines, and the limitations associated with DataLogr. Recently, another area code was added to the metro area, and an affected organization with 100 DataLogr metadata files would need to enter each file to update the fields that contain the phone numbers. Until these issues are resolved, it is difficult to convince many of these organizations to adopt DataLogr as their metadata tool. For this reason, many are convinced that a better metadata tool will evolve from DataLogr or another program that will provide cost effective metadata capabilities.

Recommendation:
Expand metadata to
document information about
map products and tabular
datasets.

Finally, with regards to the conceptual side of metadata, efforts are necessary to provide more attractive options and uses for metadata. Some consideration must be given to documenting map products and tabular datasets. Frequently, maps will be displayed, but no one will know which layers were used, or the steps taken to create the map. Tabular datasets must also be documented because they can add value to geo-spatial datasets. Using more graphic representation of information would make metadata more attractive too. An example would be to include a link to a status map that displays the spatial limits of a dataset.

RECOMMENDATIONS

The road to metadata implementation is filled with obstacles at every turn. There is no single strategy or approach to acquiring full compliance. Only time, diligence, and innovation will achieve many of the objectives that MetroGIS envisions for metadata implementation.

In order to address the time and diligence issues, MetroGIS must continue to support the dissemination of metadata information to more organizations in the metro area, and maintain contact with those already aware of these endeavors. This does not require a full-time commitment; however, there should not be any disruption in activities, so as to assure the local GIS community that MetroGIS is committed to metadata implementation. A disruption of support to metadata implementation may be interpreted as a passing MetroGIS fad, and some of the initial efforts will have been lost. Another risk is that organizations may develop their own documentation procedures rather than creating compliant metadata. This group could include new developers and users, or those who were introduced to the benefits of metadata through this MetroGIS' project, but prefer to develop their

own procedures.

Another part of the commitment equation requires that MetroGIS continues to offer support to organizations interested in producing metadata, but lack the resources. Support for these organizations requires a continuation of metadata workshops and interns to assist with metadata development. As mentioned, workshop attendance is necessary to assure that participating organizations and interns understand the same issues and metadata requirements. This will provide an environment that creates less disruption for the participating organization because both sides will understand the others needs.

Recruiting interns for metadata implementation remains an obstacle in itself. As past experience shows, interns are not especially dependable; therefore, recruitment needs to start early and include direct contact with the candidates. MetroGIS needs to get more directly involved with the faculty members, to assure a focus on this issue.

Geography departments in both the public and private universities provide a sufficient pool of students with adequate training in geo-spatial analysis and GIS. Posting a job announcement, and pursuing active recruitment, should lure students to the internships.

The method used to draw students to the internships will depend on the types of programs each program has in place. Macalester College students in the geography department must perform 140 hours of service during the course of one 15-week semester to fulfill part of the graduation requirement. The Department of Geography at University of St. Thomas also offers internship opportunities to its students; however, it is not compulsory. The Department of Geography and the Center for Urban and Regional Studies (CURA) at the University of Minnesota provide internships for credit, though student compensation must be taken into account. A cost sharing agreement for metadata service could be arranged between interested organizations and MetroGIS, with MetroGIS taking the lead in hiring and training the interns, and coordinating their activities.

The Minnesota GIS/LIS newsletter and other regional GIS-related media should be used to acknowledge achievements and provide activity updates. Carver County appears to have taken the initiative towards metadata compliance, and it should be commended for its efforts. The collaborative effort between MetroGIS and Washington County deserves to be mentioned as well. MetroGIS should arrange for an short awards presentation at the annual GIS/LIS conference.

While all these efforts are in motion, MetroGIS must target strategic contacts for metadata implementation, especially at organizations facing staff transitions, starting new GIS operations, and expanding or creating new departments. At some point, MetroGIS may choose also to return to organizations where contacts had rejected metadata during initial meetings. These situations may require that an administrator is contacted rather than someone in a GIS department, because the administrator can recognize the importance of metadata in light of a key staff person's departure who hasn't established any dataset documentation procedures. The importance of contacting the right person at an organization cannot be overstated, and provides that link to assuring that metadata development is written into job descriptions and performance reviews, and is written in contracts for data development and data sharing agreements. The administrator must also understand that time taken away from other GIS efforts to produce metadata, is time well spent for both individuals in the department and the organization as a whole.

Another important link in the effort involves cooperation with the private profit and non-profit companies in the metro area. Most of these organizations have been producing metadata, or some form of documentation, so assurances must be made that they produce compliant metadata for themselves and their clients. In many cases, these clients represent small communities that have no in- house, or small GIS operations. In both situations, the private firm can produce metadata for the datasets as part of the contractual agreement with the client. The firm can also offer to assist with producing metadata for datasets when public organizations have produced their own datasets internally. The private sector provides experience and valuable resources for targeting smaller public entities, which in turn allows MetroGIS to focus on the large data developers and users in the metro area.

Finally, data developers and users need assurances that improvements will be made in the development of an adequate metadata software tool, that complements or enhances DataLogr's capabilities. Universal changes to DataLogr files must be made one file at a time. DataLogr must function in a manner that allows for universal changes in multiple files. There have been some comments made about the inability to change DataLogr's edit window size. A spell check option has been recommended too. Providing the capabilities to attach additional graphics, index maps, and links to data catalogues are some of the other features that could make DataLogr a more attractive package for metadata producers.

Some professionals have mentioned the use of data catalogues as part of the solution to reaching full metadata compliance. The argument is that the data catalogue could be upgraded to include all the necessary information to make it a fully compliant metadata file. Information in the data catalogue would serve the internal needs of an organizations, but any request for a dataset would mean that the fully compliant metadata file would need to be completed and attached to the dataset.

CONCLUSIONS

There are no easy solutions to metadata implementation. It will take a combination of short and long-term strategies involving cooperative efforts, incentives, and new ideas to make metadata part of mainstream activities in the GIS community.

Currently, it appears that most of the efforts in the realm of metadata are focused on the conceptual side of metadata rather than the reality of implementation. While metadata are about sharing dataset information, more information needs to be shared about metadata. A discourse is needed between organizations engaged in metadata implementation to learn of their experiences, and gain insight from the success stories and failures that have been experienced with their endeavors. Sharing experience is critical to successful data sharing.

These efforts will move the discussion away from the conceptual facets of metadata and direct metadata implementation experiences toward larger audiences in the GIS community. The beer making practices of the Mesopotamian people, oral tradition maintained this art for centuries within the society. As the idea spread beyond Mesopotamia, the ingredients and practices changed over time as beer brewing reached many diverse cultures and growing populations around the world. As the demand for beer recipes grew, the written medium quickly passed the oral tradition because information could be disseminated more quickly and accurately to breweries across the continents. There would be no need to travel to Baghdad to enjoy a good beer.

REFERENCES

Callahan, Shawn, David Johnson, and Paul Shelley. 1996. "Dataset Publishing- A Means to Motivate Metadata Entry," in the First IEEE Metadata Conference Proceedings, Silver Spring, Maryland, 16-18 April, 1996.

Lynne Bly and Associates. 1997. "Participant Roles and Responsibilities White Paper," Prepared for the Metropolitan Council Data Finder Web Site Project.

 Murray, H.A., 1938. Explorations in Personality. New York: Oxford University Press.