

## Organizing geographic information: The creation of the National Spatial Data Infrastructure

Bradley Wade Bishop, doctoral student, Florida State University, 101 Shores Building, Tallahassee, Florida 32306-2100; Phone: 850-645-7281, Fax: 850-644-9763 email: [bwb06c@garnet.acns.fsu.edu](mailto:bwb06c@garnet.acns.fsu.edu).

### Introduction

Prior to government regulations issued in OMB Circular A-16 in 1990, the organization and dissemination of geospatial data collected by the United State's governments were unregulated and informal. Circular A-16 called for the creation of the Federal Geographic Data Committee (FGDC) and charged the FGDC with creating and implementing metadata standards for geospatial data. The importance of geospatial interoperability and metadata standardization between agencies was amplified in 1994 with EO 12906. EO 12906 called for the creation of a National Spatial Data Infrastructure (NSDI). A Spatial Data Infrastructure (SDI) is the coordination of the creation, collection, dissemination, and storage of spatial data between stakeholders in a spatial data community (Williamson, Rajabifard, and Binns 2005). This study explores the choices made by the FGDC on how to classify geospatial data across agencies and attempts to assess compliance among federal, state, and local governments to metadata standards in the creation of the NSDI.

### Background of the National Spatial Data Infrastructure

The National Spatial Data Infrastructure's goal is to reduce the duplication of work among agencies and increase interoperability between agencies. These goals allow for more formalized dissemination and storage of geospatial data. To highlight the extent of this undertaking, any agency that collects addresses or any place-bound information is an agency that creates geospatial data. Eighty thousand agencies create geospatial data and in 2005 it was estimated that this national capital asset cost four billion dollars a year to create and manage (Masser 2005).

Despite the cost and complexity of the NSDI's task to produce a National Map that would contain all datasets collected by all government agencies, the job is unthinkable without accepted metadata standards (Masser 1998). By populating one source with a standard format each agency could more easily use other

agencies' data as opposed to the costly operation of duplicating the same spatial data or misuse of data due to inadequate metadata. Some examples of geospatial data that agencies could share are street layers, topography, and county boundaries. This would allow agencies to work more closely together on issues that impact an entire community or region.

The organization of the United States government complicates the creation of the NSDI. State and local governments have the responsibility to collect cadastral data for land title records and land taxation purposes (Masser 1998). Federal government agencies and other departments, however, collect several types of geospatial data on various levels. The United States Geological Survey (USGS) and the United States Bureau of the Census are examples of two large stakeholders in the creation of the NSDI. The USGS classifies public lands and the Census Bureau collects information of both the distribution and the attributes of the country's population. The coordination and buy-in of all such stakeholders to metadata standardization was the responsibility of the FGDC. The FGDC also needed to create a metadata standard that was flexible enough for use by a variety of agencies with diverse types of current and future data.

Geographic information professionals at every agency that collect geospatial data are also stakeholders in this buy-in and the accuracy and interoperability of their agencies' data depends on the accompanying metadata. The proper use and storage of geospatial data is crucial to the stewardship of the environment, national defense, and the economic growth of the country. Ultimately, the management of environmental resources and economic development of the United States affects every person on the planet. This does not overstate the importance of metadata as the smallest building block in the nation's land use management.

In addition to the coordination and standardization of spatial data collected by government agencies, the FGDC not only needed to formulate a process to disseminate the geographic information between agencies, but to end-users in the public. The document that established the FGDC and also summarized their duties was titled *Coordination of Surveying, Mapping, and Related Spatial Data Activities*, where the focus was interagency compliance for interoperability (OMB 1990). EO 12906 strengthened the policies described in Circular A-16 and raised the visibility of the FGDC by creating the NSDI and charging them with the development of GIS clearinghouses accessible to the public (Clinton 1994). The FGDC continues

as an interagency committee addressing issues related to the technical implementation of the NSDI, but also the dissemination of data beyond the agencies (OMB 2002).

The efforts of the FGDC to create a clearinghouse of all government geospatial data will allow more of the public to enter debates of land use and other spatial issues, e.g. gerrymandering, public facility location, etc. However, for now a user needs a good understanding of geospatial metadata and the organization of government agencies that produce different types of data to locate useful information. Currently the vision of the NSDI and the metadata standards of the FGDC are designed for GIS professionals and therefore limit findability beyond the profession.

## **Geospatial Data**

Geospatial data adds the dimension of location to all defense, economic, social, and environmental management. To optimize the impact of government programs, knowledge of the spatial aspects of policy issues is crucial. Whether the issue is protecting wetlands or choosing the site of a homeless shelter, geography matters. Government agencies understand the value of geography, but the value of interoperability between agencies was not addressed in the United States until the creation of policy for the NSDI and its implementation by the FGDC.

In the most recent global survey, sixty-seven countries had implemented a Spatial Data Infrastructure and thirteen had projects for implementation (Cromptoets, Rajabifard, and Williamson 2004). Geospatial data or geographic information are data that define the geographic location of features and boundaries on Earth. In 1994, the United States became the first country to start implementation of a SDI. In 2002, revised Circular A-16 added homeland security as another important aspect for the creation of functional NSDI.

## **Clearinghouse and Metadata Standards**

The enormity of the FGDC's task to coordinate the standardization and dissemination of geographic information is evident in Geospatial One Stop's (GOS) GIS for the Nation poster developed by the Environmental Sciences Research Institute (ESRI). GOS is the clearinghouse for the nation's geospatial data. A clearinghouse is a searchable database of all geographic information's metadata that provides links

to agencies that produce and disseminate geographic information (Masser 2005). The GIS for the Nation poster separates the spatial data of the nation into themes, scales of use, guidelines and datasets associated with the each type of geographic information (ESRI 2007).

Examples of data themes collected by agencies include transportation, hydrography, emergency operations, and satellite imagery. The scales of use range from federal to neighborhood level data. FGDC guidelines describe the accuracy required for proper use of the data and the frequency of updating the data. Datasets associated with different types of geographic information are also provided to assist agencies in determining the datasets required for production for each type of geographic information. Datasets can be coverages, layers, geodatabases, shapefiles, Triangulated Irregular Networks (TINs), Light Detection and Ranging (LIDAR), etc. The table below illustrates the organization with a simple type of geospatial data: cadastral.

Data Theme	Scales of use	Guidelines	Geographic Datasets
Cadastral: tax parcels, includes zoning, rights, easements, etc.  Used for tax mapping and surface ownership maps	Local	Collected at neighborhood level using survey techniques	Table of address and theme attributes
	Tribal		
	State	At the neighborhood level the parcel is the primary mapping unit	Feature class of address points
	Federal		

Besides the complexity of organization, the table also illustrates the importance of metadata. The key to the success of the NSDI is the adoption of metadata standards by agencies at every scale of use. Metadata is supporting information for geospatial data that describes the quality, type, attributes, and limitations of the data. Without the added value of metadata, geospatial data is of little use beyond the persons or agency that originally created the data. FGDC guidelines for metadata not only save time in sharing of the data, but also in its creation. Accuracy for elevation when determining the extent of flood planes requires greater precision than the point location of a hospital to assess its service area. The limitations of different types of data are required in the metadata and help GIS professionals recall this during their creation.

## Organization and collection

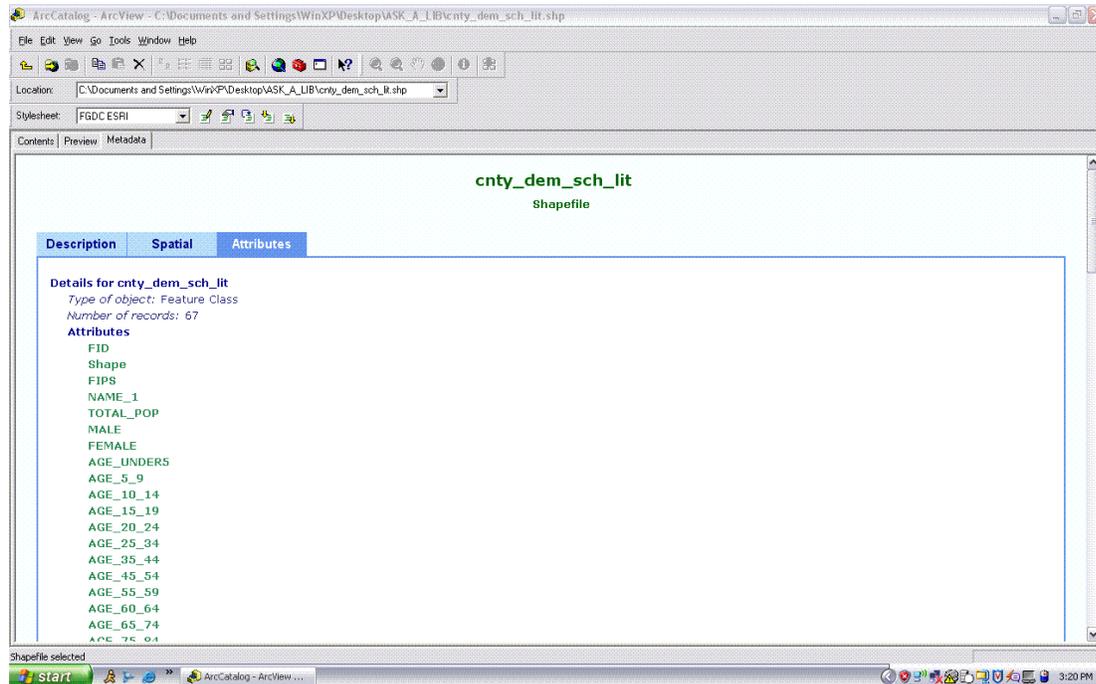
The organization and collection of geospatial data includes the adoption of FGDC metadata standards and the standardization of the workflow of all information within agencies. Geographic information professionals do not receive the training to implement information management. Geographic information management contains unique information problems.

A person updating geospatial data in the field can instantly save this data collected on a handheld global positioning (GPS) device to an authority file dataset on a server at the agency. The complications related to this simple task illustrate the complexity of the information problems in geographic information. There are issues related to accuracy of the metadata and versioning.

Advancements in the technology of both handheld GPS devices and remote sensing has increased the accuracy of the geospatial data collected. The newest data are the most accurate -- but it is still not fiscally feasible to field-test the entire country. Which version of the data should the government use? Who has the authority to alter existing spatial data? Data collected prior to the adoption of metadata standards retains metadata that are inoperable with new metadata and in many cases lacks metadata that would increase its value and prevent misuse. Examples of valuable metadata are the time it was collected and how it was collected. GPS and remote sensing devices can timestamp the moment the new data was collected, but these methods still require input from the user for other metadata.

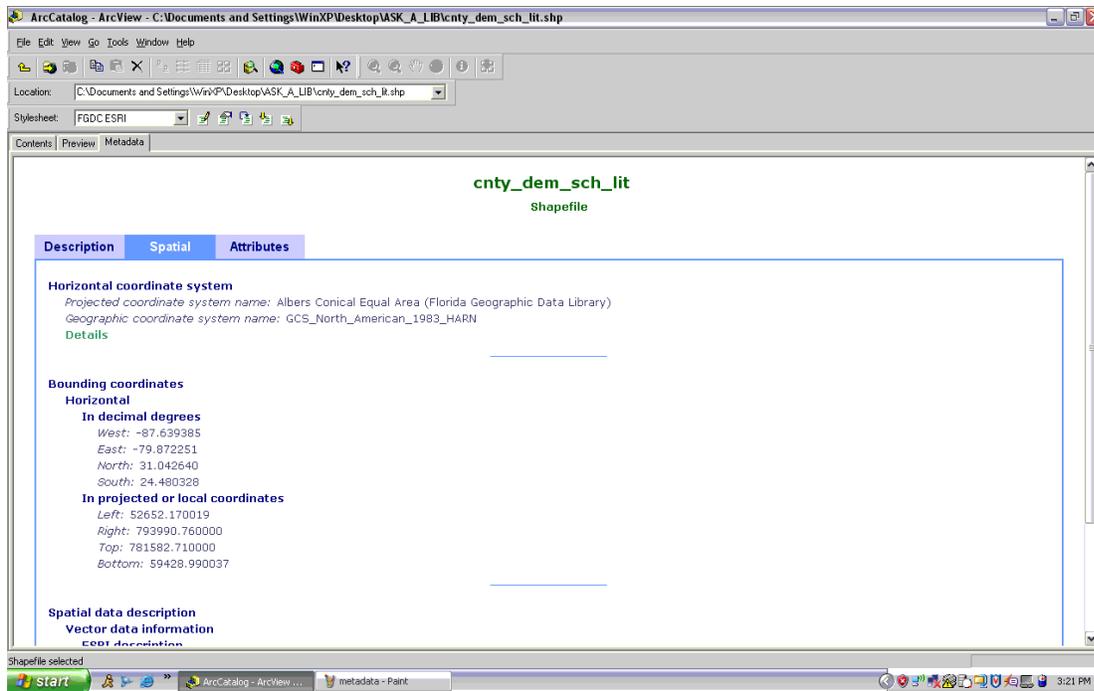
Along with these information management issues, the FGDC must also forge partnerships across agencies (Harvey and Tulloch 2006). Each agency will have different solutions to unique information management issues, but the products must retain enough metadata so other agencies can use the data appropriately and understand the limitations of the data. Someone in marine biology may not be knowledgeable enough about soil science to know the difference between sandy loam and clay, but the metadata should be detailed enough to provide a researcher with information to see the difference in the data. Below is a screen shot of ESRI's ArcCatalog. This illustrates the importance of metadata for users of geographic information because a user cannot tell what it is or how it could be used without reading the metadata.

Each piece of metadata has a description once you open up an attribute, if the creator of the metadata provided one.



Geospatial metadata is more crucial for use of geographic data than attribute metadata. One example of geospatial metadata is a map projection. A map projection is any of many methods used in cartography to represent the two-dimensional curved surface of the earth or other body on a plane (Wade and Sommer 2006). Data will not appear in the correct place if two different data sets are projected in different map projections. Reprojection of data further distorts the accuracy of the original data and can lead to unfortunate results. Therefore, users need to know the original projection of the geospatial data prior to using it.

A screen shot below of ESRI's ArcCatalog illustrates the importance of geospatial metadata for users of geographic information, because without spatial metadata the data cannot be properly used. The geospatial data includes bounding coordinates of the data's extent. The bounding coordinates are the northernmost, southernmost, easternmost, and westernmost points of the data (Hill 2006). The FGDC also accepts g-ring geospatial data, which is a string of points that forms a line around the data's extent. A g-ring is most often used for data that do not fit into a quadrangle, e.g. river.



## Response from information professionals

The preeminent map librarian, Mary Larsgaard, advises libraries to follow established metadata standards such as ISO 19115 or FGDC as opposed to creating their own (Larsgaard 2005). Small GIS shops and research facilities are as important to the creation of the NSDI as the Census Bureau. The FGDC metadata standards resemble MARC records, with fields for author, publication date, edition, publication time, series information, and others. The unique fields are geospatial data presentation form and georeferencing data. These include the theme, type of dataset, and the extent of the data's coverage. Some of these correspond to fields in the MARC record and allow for crosswalks between the two. The major goals of the FGDC metadata standards are listed below:

- maintain the organization's investment
- provide a catalog of the organization's holdings
- allow external users to process and interpret the data (Hill 2006).

These goals resemble the purpose of any catalog. Unfortunately, few geographic information professionals receive information management or cataloging experience in Geography or Urban Planning courses.

## GIS education

Unlike other professionals, geographic information professionals lack a standardized curricula or tests for competency. This is in part due to relatively recent development of GISystem technologies used in many fields and the federal policies to increase the management and dissemination of geospatial data. Imagine medical doctors and lawyers practicing without licenses in mass. The GIS Certificate Institute exists to attempt to quantify the necessary skill sets required to be a geographic information systems professional (<http://www.gisci.org>). The education portion of the certificate does not discern between degree programs and non-degree programs or the varying courses taught in both.

A brief review by this researcher of ten GIS degree programs in the southeast found no courses offering content in information management. Without information management training included in GIS curricula, the spatial community will continue to exist with well-trained geographers with novice information skills. The FGDC offers training on its metadata standards, but is it effective?

## Methodology

To assess the understanding of FGDC metadata standards among geographic information professionals and the success of their implementation, the researcher identified geographic information professionals with metadata duties and created a semi-structured interview. The interviewer identified these geographic information professionals based on a recommendation from the staff at the Florida Resources and Environmental Analysis Center (FREAC). The single contact led to a snowball sample based on recommended professionals from the first contact. For further study, the researcher will expand the variety of professionals interviewed.

The questions are intended to explore metadata geographic information professionals' perspective on metadata and the challenges they have faced implementing metadata standards and how they have helped other geographic information professionals understand the value added of metadata. After the interviews, the researcher reviewed the answers for trends.

## Interview protocol

1. What is metadata?
2. What challenges have you faced implementing metadata standards?
3. How have you helped others understand the value added of metadata?

## Findings

Perhaps due to the small sample, the collected experiences and answers were similar. All four interviewees had a good understanding of metadata. "Metadata is the meaning of data" (Santi 2007). Jill Truby, the Metadata Librarian/Coordinator of the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute, corrected the researcher's first question, "you must mean, what are metadata? Metadata is plural" (2007). An understanding of metadata was apparent among the geographic information professionals whose job duties required it, but how did these few perform outreach into their organizations?

The challenges faced in implementing the metadata standards in Truby's case included the considerable time it took to convince upper management to require workers compliance with the standards. Truby is a trained trainer for the FGDC on the metadata standards and assisted her organization in complying, but estimated the process took ten years. It helped metadata compliance when grants began requiring corresponding metadata from researchers for their geospatial data. The researchers also found metadata made geospatial metadata easier to retrieve and the researchers were more supportive of the standards.

Stephen Hodge, Florida Resources and Environmental Analysis Center (FREAC) geographic information professional, provided some insight into the system's method of dissemination before NSDI policy.

"Persons would call me looking for spatial data. Sometimes, we [FREAC] would have it. Sometimes, I would know who they should call. Sometimes, I would know who might know who had it. It was a very time consuming process that took away from my real work (Hodge 2007)."

Metadata standardization has made the search for geospatial data much easier. Hodge also faced challenges when implementing the metadata standards, but geospatial metadata management tools have been developed that do not let persons publish any geospatial data without providing the minimum record of corresponding metadata. The user can select which standard they are complying with, but FGDC is the most used at FREAC. Christi Santi, the GIS Librarian at the Fish and Wildlife Research Institute, has

experienced issues with interoperability between multiple metadata software that is used by researchers on different projects.

To help others understand the value added of metadata, the interviewees have helped end-users by showing the speed of locating data with metadata. Although only half of the small sample had a LIS/IS education, the knowledge of metadata and its value are apparent to all geographic information professionals.

## Conclusion

The process to develop and implement geospatial metadata standards began in 1990 with the creation of the FGDC. Seventeen years later, after more policy and planning, the progress toward standards and a functional NSDI seem promising based on this small sample. Unfortunately, those geographic information professionals interviewed were in research settings that depend on grant funding to operate. By reviewing online sources for geospatial data, the researcher knows the accessibility to government geospatial data is not always as ideally disseminated as the data in research settings. The sampling method used led to this. Further research of this assessment must include geographic information professionals that do not operate in research settings.

Access to local government information varies greatly between city, county, and state governments. The issues related to geospatial data are just one portion of e-government and perhaps not governments' first priority when disseminating information electronically. Luckily, the success of metadata standardization makes the dissemination of geospatial data possible in formats other than through the Internet. The next step for the FGDC is encouraging all agencies to populate clearinghouses, such as Geospatial One-Stop, with geospatial data with accompanying metadata. The FGDC will also need to continue to update metadata standards for specific fields and different types of geospatial data as well as comply with ISO standards. This compliance will facilitate the sharing of geospatial data between countries, because environmental stewardship and economic development are not concerns unique to the United States.

## References

- Clinton, William. "Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure." Executive Order 12906, April 13 1994, pt. 6. <[http://frwebgate2.access.gpo.gov/cgi-bin/waisgate.cgi?WAI\\_SdocID=046113206402+0+0+0&WAI\\_Saction=retrieve](http://frwebgate2.access.gpo.gov/cgi-bin/waisgate.cgi?WAI_SdocID=046113206402+0+0+0&WAI_Saction=retrieve)> (5 April 2007).
- Crompvoets, A. B., Rajabifard, A, & Williamson, I. "Assessing the worldwide developments of national spatial data clearinghouses." *International Journal of Geographical Information Science* 7, no. 18 (2007): 665-689.
- Environmental Systems Research Institute. "GIS data for the nation." June 2007. <<http://support.esri.com/index.cfm?fa=downloads.dataModels.filteredGateway&dmid=42>> (8 October 2007).
- FGDC. "Geospatial Metadata." 2005. <<http://www.fgdc.gov/library/factsheets/documents/metafact.pdf>> (20 March 2007).
- FGDC. "Policy statements for federal government geographic data sharing." 1992. <<http://www.fgdc.gov/library/factsheets/documents/datashare.pdf>> (20 April 2007).
- Harvey, F., and D. Tulloch. "Local-government data sharing: Evaluating the foundations of spatial data infrastructures." *International Journal of Geographical Information Science*, 20, no. 7 (2006): 743-768.
- Hill, Linda. *Georeferencing*. London: MIT Press, 2006.
- Hodge, Stephen. Conversation with author , 10 June 2007.
- Larsgaard, Mary. L. (2005). Metaloging of digital geospatial data. *The Cartographic Journal*, 42(3), 231-237.
- Masser, Ian. *Governments and geographic information*. London ; Bristol, PA: Taylor & Francis, 1998.
- Masser, Ian. *GIS worlds : Creating spatial data infrastructures* . Redlands, CA: ESRI Press, 2005.
- Ramdeen, Sarah. Telephone conversation with author, 2 October 2007.
- Santi, Christi. Telephone conversation with author, 31 July 2007.
- Truby, Jill. Telephone conversation with author, July 2007.
- United States. Office of Management and Budget. "Coordination of Surveying, Mapping, and Related Spatial Data Activities." 1990. <[http://www.whitehouse.gov/omb/circulars/a016/a016\\_rev.html](http://www.whitehouse.gov/omb/circulars/a016/a016_rev.html)> (22 March 2007).
- United States. Office of Management and Budget. "Coordination of geographic information and related spatial data activities." 2002. <<http://purl.access.gpo.gov/GPO/LPS45567>; <http://purl.access.gpo.gov/GPO/LPS45567>> (8 April 2007).

Wade, Tasha, and Shelly Sommer, eds.. *A to Z GIS : An illustrated dictionary of geographic information systems*. Redlands, Calif.: ESRI Press, 2006.

Williamson, I.;A. Rajabifard; and A. Binns. "Challenges and issues for SDI development." *International Journal of Spatial Data Infrastructures Research* 1, (2006): 24-35.