

Enabling FGDC Metadata Development in Biodiversity and Ecosystem Sciences

August 15, 2005

Eric Landis
Natural Resources Information Management
Stillwater, MN
(651) 275-0775
elandis@ix.netcom.com

Table of Contents

1. Executive Summary	3
2. Introduction	5
2.1 Project Objectives	5
2.2 Study Process.....	5
3. Summary of Findings	6
4. Findings and Discussion	7
4.1 Organizational Issues	7
4.1.1 <i>Management</i>	7
4.1.2 <i>The mission statement, incentives and policies</i>	8
4.1.3 <i>The organizational structure and personnel</i>	9
4.2 The process.....	10
4.3 Metadata Education	12
4.3.1 <i>Metadata Training</i>	12
4.3.2 <i>Higher Education</i>	13
4.4 Metadata Authoring and Presentation Tools	14
4.4.1 <i>Authoring Tools</i>	14
4.4.2 <i>Presentation Tools</i>	15
4.5 The FGDC Standard.....	16
4.5.1 <i>Element Definitions</i>	16
4.5.2 <i>First Impression</i>	16
4.5.3 <i>Controlled Vocabulary</i>	16
4.5.4 <i>Transition to ISO</i>	17
5. Case Studies: Successful Metadata Development	17
5.1 Florida Fish and Wildlife Research Institute	17
5.2 USGS Forest and Rangeland Ecosystem Science Center	18
5.3 St. Croix National Scenic Waterway.....	19
6. Recommended Strategies for Successful Metadata Development within Natural Resource Management Units	20
6.1 Planning and Process	20
6.1.1 <i>Develop an Information or Data Management Plan</i>	20
6.1.2 <i>Create policies and directives to enforce metadata development.</i>	20
6.1.3 <i>Include data and information interests within the organization’s mission statement, goals and objectives.</i>	21
6.1.4 <i>Re-define the organizational structure and personnel to facilitate metadata.</i> ..	21
6.1.5 <i>Establish a metadata development process that corresponds to the project life cycle</i>	22
6.1.6 <i>Adopt existing recognized controlled vocabularies.</i>	23
6.2 Management Support	24
6.2.1 <i>Educating managers about metadata</i>	24
6.3 Promotion and education	24
6.3.1 <i>Capitalize on FGDC metadata training opportunities</i>	24
6.3.2 <i>Develop agency or organization metadata support websites</i>	25
6.3.3 <i>Promoting metadata development with recognition and incentives</i>	25
6.4 Evaluating and Prioritizing Recommendations – An Example	25
7. Recommendations to FGDC	27
7.1 Educational.....	27
7.2 Outreach.....	27
7.3 Tools.....	28
7.4 The Metadata Development Process	29
7.5 The Standard.....	29
7.6 Other Recommendations	29
Annex A: FRESC Electronic Metadata Interview	30

1. Executive Summary

This document reports on a series of interviews, meetings, case studies, workshops and observations to discover common obstacles to developing Federal Geographic Data Committee (FGDC) metadata and offer recommended strategies for increasing the frequency and improving the quality of FGDC metadata records. The study is based on input regarding FGDC's Content Standard for Digital Geospatial Metadata (CSDGM) - version 2. The majority of findings and recommended strategies, however, are applicable to the development of metadata under all standards.

The pre-study assumptions for reasons that metadata is not developed emphasized a lack of time and expertise, issues of data ownership and security, rigorous standards, and difficult to use authoring tools. While these issues came up during discussion with study participants and are reflected directly or indirectly in this report, other critical issues were identified. The most frequently cited obstacles to developing metadata were:

1. a lack of management support and understanding of the value of, and process for, metadata development
2. an absence of wide-ranging education and promotion of the benefits of, and process for, metadata development
3. the development of metadata authoring and management procedures that correspond with project lifecycles from project design to data archiving and re-use
4. a need for an organizational structure and personnel within resource management units to support metadata development and management
5. a lack of easy-to-use interactive authoring tools designed for occasional metadata developers
6. unclear definitions for the FGDC metadata elements

Other hurdles to metadata development were identified by participants and are noted in Section 3 and summarized in Section 5 of this report. While many of these obstacles need to be addressed by resource management units such as National Parks or Ranger District offices, some require national attention that is best addressed by the FGDC.

Recommended strategies to resource management units focus on recognizing and acting upon the role that data and metadata have within the enterprise. Ten recommendations are made:

1. Develop planning documents for information and data management activities.
2. Implement policies and directives to enforce metadata development.
3. Incorporate information management and provision into the organization's mission statement, goals and objectives.
4. Develop the organizational structure and assign personnel to support information management, much as is done for financial management.
5. Establish a metadata development and management cycle that corresponds to project lifecycles.
6. Adopt existing recognized controlled vocabularies.
7. Educate managers on the benefits and process of metadata management.
8. Capitalize on FGDC metadata training opportunities.
9. Develop agency or organization metadata support websites.
10. Promote metadata development with recognition and incentives.

Recommendations to FGDC (Section 7) focus on education and promotion of metadata, and providing implementation support. Recommendations include:

1. Expanding the Cooperative Agreement Program (CAP) to include additional categories for supporting metadata development.
2. Establish metadata promotion campaigns for a broad, multi-level community from data owners to agency administrators.
3. Establish educational programs for non-metadata developers, including managers, administrators, and agency heads.
4. Facilitate the development of easy-to-use, interactive metadata authoring tools for occasional and beginning metadata developers.
5. Clarify, with real-world examples, the CSDGM element definitions.
6. Provide support for metadata implementation activities including the development of generic business and management plans for metadata development, and sample data management policies and templates.

2. Introduction

The focus of this project is to explore FGDC metadata development in national parks, ranger districts, tribal groups and other federal, state and regional management units that maintain data collection and archiving activities for the biological and ecological sciences. This study is based on findings from interviews, workshops, case studies, and observation. It identifies obstacles to, and strategies for, the successful implementation of FGDC metadata development.

2.1 Study Objectives

The objectives of this project are to:

1. Review and document current metadata policies, procedures and practices across a representative sample of natural resource agencies at the federal, state and regional jurisdictions.
2. Document inhibiting and facilitating aspects of metadata development when utilizing the FGDC metadata standard.
3. Develop recommendations for improving the frequency, quality, utility and distribution of FGDC metadata for biological and ecological datasets.

2.2 Study Process

This study was conducted from August 2004 through February 2005. In April and May of 2004 a preliminary study was conducted as a follow-up to a FGDC CAP metadata training workshop held in Ashland, Wisconsin. Relevant findings from that study have been incorporated into this document.

Approximately 40 individuals from the following groups participated in the study. Some were interviewed in person or by phone, others participated in metadata development observation and some participated through regularly scheduled meetings and workshops.

- Metadata trainers
- Metadata developers
- Resource managers
- FGDC administrators
- Metadata creation tool developers
- Data managers and archivists
- Educators
- Clearinghouse managers

Target questions for each group were established (see earlier document, "Enabling Metadata – Step 2"), but most discussion was free flowing with the questions utilized only as a guide and for assurance that all subject areas were considered.

In addition to direct interviews and observations, the investigator attended the following workshops and meetings that included discussion relevant to this study's objectives:

- National Conference on Digital Government Research, May 18-21, 2003, Boston, MA, Birds of a Feather Session - *Developing an Integration Infrastructure for the Statistical Knowledge Network: Metadata and Standards*
<http://www.dgrc.org/dgo2003/START.html>

- National Conference on Digital Government Research, May 24-26, 2004, Seattle, WA, Birds of a Feather Session – *Ecoinformatics*
<http://www.diggov.org/library/library/dgo2004/>
- FGDC 2004 CAP Program Kick-Off Meeting, Sept. 28-29, 2004, Denver, CO
- Workshop on Biodiversity and Ecosystem Informatics, Dec. 13-15, 2004, Olympia, WA. <http://www.evergreen.edu/bdei/>

3. Summary of Key Findings

The following is a compiled list of the key findings identified in Section 4 and has been used to develop recommended strategies for successful metadata development ([Section 6](#)).

Organizational Issues

1. Managers require more knowledge of the need, process, cost and benefit of metadata development.
2. Currently, metadata programs are principally initiated by data handlers, not managers and administrators.
3. Incorporating information management into enterprise-wide mission statements has facilitated an understanding of the need for good data management practices.
4. Maintaining policies and incentives for good data management practices has assisted organizations carry-out metadata development.
5. Successful metadata development programs maintain personnel specifically assigned to oversee, conduct and/or assist with data management activities.
6. Successful data management personnel are in a highly recognizable position for the entire organization.

Metadata Development Process

7. Metadata development programs that rely on a single individual seldom succeed.
8. Developing metadata in a series of process steps conducted by multiple individuals was strongly encouraged by successful metadata developers.
9. Participants reported that initiating metadata development during the project planning phase increased the likelihood of success.

Education and Training

10. FGDC's metadata training program is well-received.
11. Trainees, in general, feel a need to expand the training in scope and duration.
12. On-site (mobile) training has the added benefit of "broadcasting" the value and requirements to the trainees' co-workers, including their supervisors.
13. Follow-up activities or monitoring after metadata training workshops is reported to improve metadata learning and success rates.
14. Those who have been in the position of searching for and accessing data via FGDC metadata records are more likely to be successful metadata developers.
15. Higher education institutions seldom incorporate metadata development into current resource or data management curriculums.
16. Higher education institutions are willing to explore how metadata can be incorporated into resource or data management curriculums.
17. Some higher education administrators may need to be convinced of the value of including metadata development in course work.

18. No standard metadata curriculum or course materials are readily available to college and university instructors.

Metadata Tools

19. Currently available metadata authoring tools are complex for most biologists, researchers and others being asked to develop metadata records.
20. Beginning and novice metadata developers are seeking easy-to-use, open source tools that require little or no training.
21. Metadata developers feel that improving the visual appeal and usability of the end product will improve the willingness and frequency of developing metadata records.

The FGDC Standard

22. Definitions for FGDC elements are vague for even experienced metadata developers.
23. FGDC standard appears complex and intimidating on first introduction.
24. Some requested that the standard be organized by users' needs.
25. The lack of readily available controlled vocabularies contributes to mis-cataloging of datasets.
26. There is little information available regarding the planned transition to the ISO metadata, causing some anxiety.

4. Discussion of Key Findings

The findings of this study are organized by issues related to the organization, the metadata development process, education and training, metadata authoring and presentation software tools, and the FGDC metadata standard.

4.1 Organizational Issues

4.1.1 Management

One of the most cited hurdles to metadata development within an organization is the fact that managers (those that supervise GIS specialists and biologists) are not fully informed of the benefits of metadata or what is involved in producing metadata records. One study participant noted, "The basic problem is that managers in agencies don't believe in the value of metadata. They haven't had to ensure or warranty data quality." Another participant stated, "The reality of what it takes to do a good job of documentation is lost on many managers."

The organizations that maintained successful metadata programs were those whose managers openly supported metadata management. Their support was evidenced by their willingness to allocate resources to metadata development and broader data management efforts. For example, their employees were more likely to attend training, have the necessary data management tools, and have time and funding allocated to data management activities. The managers' support originated from either their attendance at a metadata workshop or strong data management values stemming from work experiences.

No reports of a manager being influenced by national initiatives, directives from agency headquarters, or any "higher" authority were made. In fact, one organization when asked how FGDC could help them, stated that educating their agencies national office personnel on the importance of metadata would be the greatest help. They observed

that getting funding from their Washington D.C. headquarters for data management activities is generally a “wild card” and noted that as data management takes place in an office setting as opposed to the field, its’ value is less recognized during the budgetary process.

While some “top-down” influences to develop metadata exist, for most situations the metadata *cause* principally relies on the efforts of biologists, GIS specialists, data archivists, and other data “handlers” to influence and educate their managers. Management support is a “must have” for organizations looking to maintain good data management practices, including metadata development.

Key findings:

1. Managers require more knowledge of the need, process, cost and benefit of metadata development.
2. Currently, metadata programs are principally initiated by data handlers, not managers and administrators.

4.1.2 The mission statement, incentives and policies

Too infrequently are the terms “data” or “information” found within the mission statements, goals or objectives of resource management organizations or units. Where they are present, a healthy information management program generally exists. The following statements are provided as examples of information-oriented goals and objectives.

“Through effective research and technical knowledge, we provide timely information and guidance to protect, conserve, and manage Florida’s marine and coastal resources.”
Florida Fish and Wildlife Research Institute

“To make PNW information most useful and accessible to our customers.” USDA Forest Service PNW Research Station

“Significant investments are made not only in production of scientific publications, but also in information and technology transfer to managers, policy-makers, educators, and the general public.” USGS Forest and Rangeland Ecosystem Science Center.

In addition to utilizing mission goals and objectives such as the above statements, some organizations rely on incentives, such as funding, to encourage metadata development. For example, the National Park Service’s Midwest Regional Office GIS program evaluation for future program funding is based in part on the existence (or absence) of metadata for existing GIS datasets. The funding evaluation guidelines state:

“Data documentation, or metadata, is critical to protecting significant investments already made in existing/ongoing GIS programs. Has the requested park demonstrated a capability to collect, maintain, and distribute metadata?” The criteria for this evaluating this activity states, *“Requesting park has an installed GIS system and has demonstrated both a willingness and capability to document and distribute metadata for existing data layers. Available documentation adheres to either legacy or FGDC-compliant standards. Information is freely available on a publicly accessible GIS clearinghouse.”*

<http://www.nature.nps.gov/im/units/glkn/>

Some organizations have instituted policies to facilitate metadata development. One such policy can be found in the USGS National Wetlands Research Center’s Policy

Issuance System. The policy, in part, states, “*Originators of scientific and other information products must initially develop metadata for all active data sets. Subsequently develop documentation for completed data sets and for historic data.*” And, “*Add metadata to the NBII Clearinghouse through the NWRC node.*” The policy also outlines the responsibilities of other personnel including the metadata coordinator, branch chiefs and center director. The data manager cited that having a visible policy that everyone had agreed to prior to collecting data assisted her in approaching researchers who had not completed metadata records for their datasets.
<http://www.nwrc.usgs.gov/>

The Coweeta Long-Term Ecological Research program established a policy “*to both increase public and private awareness of our research activities and to provide a mechanism for the appropriate transmission of information to interested parties. The data policy of the Coweeta LTER project applies to data: 1) funded by, 2) collected on-site, or 3) relevant to the LTER research program. We require metadata for all research meeting the aforementioned criteria.*” <http://www.lternet.edu/sites/cwt/>

Key findings:

1. Incorporating information management into enterprise-wide mission statements has facilitated an understanding of the need for good data management practices.
2. Maintaining policies and incentives for good data management practices has assisted organizations carry-out metadata development.

4.1.3 The organizational structure and personnel

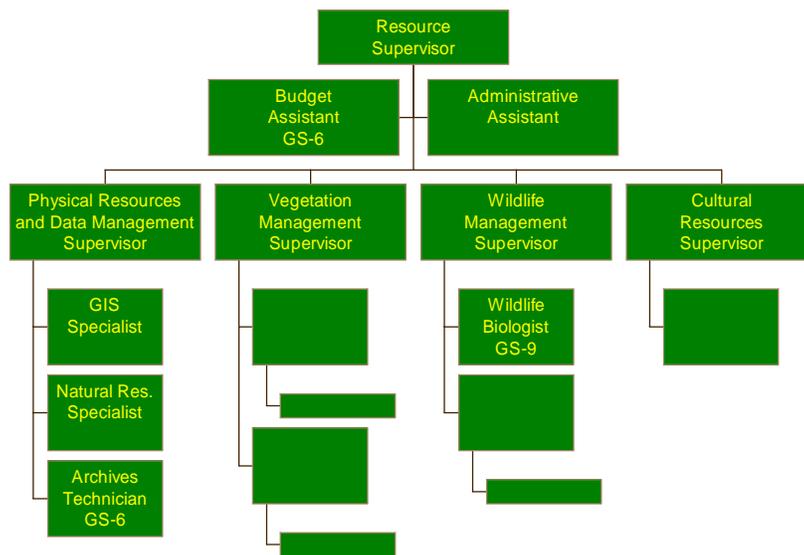
The reported procedures for developing metadata vary from organization to organization and within organizations. Metadata development by study participants follows one of three basic models, or a combination thereof – owner registration, facilitated registration or mobile registration. “Owner registration” is simply the data owner; say a biologist or GIS specialist, completing metadata records for their own datasets. “Facilitated registration” occurs when an individual other than the data owner or collector is assigned to complete the metadata record. The third model, “mobile registration” is where an assigned individual(s) travel from site to site to complete, update and/or archive metadata records.

It was recognized that assigning an individual to develop metadata, or at least monitor metadata activity, greatly facilitated the frequency and quality of metadata records. Most participants stated that a designated staff member, be it a data manager, archivist, metadata coordinator, etc., is an essential requirement for successful metadata development. “The presence of a data manager gets results,” one participant stated. Other comments included; “Don’t expect biologists to do metadata alone.” Another wrote “A metadata coordinator or editor is critical (to the effort).”

Besides assisting with the development of metadata records, *central figures* were reported to help edit records to ensure quality and compatibility within a metadata collection, arrange or conduct metadata and data management training, post metadata records, acquire and update software, create templates, and many other tasks that would “fall between the cracks” were it not for their effort.

On occasion, individuals charged with data and metadata responsibilities were not given proper authority or recognition to fully conduct their work. The organizational chart below shows one such resource management unit. While employing a data archivist to catalog, store, back-up and provide for the general management of valuable datasets most employees saw the data archivist as ancillary to their work rather than the integral part it should be. Given this perception the archivist had difficulty acquiring information to create metadata records, make back-up copies, etc. When asked how many requests are required to acquire a copy of a dataset she stated, “Many.” On the other hand, the budget assistant, when asked how many requests are required to get project financial information, she stated, “Just once.”

Table 1: Organizational chart from a natural resource management unit (Note placement of Archive Technician and Budget Assistant)



Key findings:

1. Successful metadata development programs maintain personnel specifically assigned to oversee, conduct and/or assist with data management activities.
2. Successful data management personnel are in a highly recognizable position for the entire organization.

4.2 The process

For most organizational units, developing metadata records in biological and ecosystem science has remained the responsibility of the data owner. As such, most metadata records are developed as a last step to data management responsibilities. Relying on data owners to catalogue their own datasets commonly results in a less than satisfactory frequency and quality of metadata records. These individuals are trained researchers, not information specialists. Even GIS specialists are seldom introduced to metadata in university curriculums (see [Section 4.3.2](#)). To expect these individuals to “become” information scientists in addition to their regular duties is optimistic.

It is interesting to note that prior to the advent of online data clearinghouses, metadata was developed by several individuals during the process of research, authoring, editing, publishing and cataloguing. Those metadata records usually ended up in library catalogues. Only a few study participants reported such a multi-person effort within their organization to create metadata records. The majority were “one-man shows” when it came to metadata. Serious concerns were aired in these “owner registration” systems. These concerns included time and metadata knowledge. While it was noted on several occasions that the data owner must be involved in metadata development, the success as measured by the number of records was low when it was solely the data owner that was responsible for metadata development.

Because of the complexity and importance of developing metadata, some organizations are beginning to institute a different metadata management procedure. The National Park Service is hiring data managers for networks of parks. The Florida Fish and Wildlife Research Institute maintains a metadata editor on staff to facilitate metadata development (among other duties). The Forest Rangeland and Ecosystem Science Center (FRESC) utilizes a metadata interview form ([Annex A](#)) that is passed between a metadata coordinator and the data owners. Each of these processes utilizes at least two individuals to develop and post each metadata record.

All organizations that reported successful metadata development also reported a multi-step process for creating metadata. While each organization had slightly different processes, each emphasized the importance of beginning metadata development before data is collected. One participant stated, “Ninety percent of metadata records can be done prior to data collection.” Whether the 90% figure is accurate or not, the point is that successful organizations require some metadata development prior to the commencement of data collection. Wayne (2001) identified the relationships between data development stages and FGDC metadata fields. For instance, Section 1, Identification Information, of the FGDC standard can be completed in the pre-data collection planning stages of a project. This pre-data collection metadata development step is also recognized in the FRESC metadata interview process ([Annex A](#)) where the majority of the interview questions relates to Section 1 and can be addressed during the project planning process. Florida’s Fish and Wildlife Research Institute also incorporates metadata development in their seven step “Project Life Cycle” process (see [Case Study](#)).

Table 1. Metadata fields mapped to workflow (from Wayne, 2001)

Data Development Stage	Metadata Information
Data Planning	Identification Information (Sec. 1) title, originator, abstract, purpose, keywords, content time period Data Organization (Sec. 3) point, raster, vector Spatial Referencing (Sec. 4) Coordinate system and datum Entity and Attributes (planned) (Sec. 5)
Data Processing	Data quality (Sec. 2) Completeness, positional accuracy, geoprocessing steps

Data Analysis	Data quality (Sec. 2) Attribute accuracy, analysis steps Entity and Attributes (results) (Sec. 5) Metadata Reference (Sec. 7)
----------------------	---

Key findings:

1. Metadata development programs that rely on a single individual seldom succeed.
2. Developing metadata in a series of process steps conducted by multiple individuals was strongly encouraged by successful metadata developers.
3. Participants reported that initiating metadata development during the project planning phase increased the likelihood of success.

4.3 Metadata Education

4.3.1 Metadata Training

Approximately 60% of study participants had reported that they had taken part in some aspect of FGDC’s CAP metadata training opportunity. Virtually 100% of those attendants gave positive feedback to the workshops.

In addressing how to improve the workshops themselves, three recommendations were repeated:

1. expand the training
2. conduct training sessions at the trainees site
3. conduct follow-up training activities

Most participants felt that the CAP-sponsored metadata workshops could be lengthened although there was no consensus on what material would be covered. Approximately half of those expressing an opinion for lengthened workshops felt that more hands-on metadata entry was needed. The other half felt that more introductory material such as benefits, resources, clearinghouses, tools, etc. should be covered.

Several study participants felt that metadata training should, if possible, take place at the trainees work sites. They noted that having trainers spend a day at their location would bring added benefits including;

- their supervisors seeing metadata activity and gaining an understanding of the value of metadata and what’s involved,
- trainers gaining a better understanding of the conditions surrounding their trainees daily schedule, available resources, support, tools, etc., and the added advantage of
- *forcing* trainees to “clean their datasets” before the trainer arrives.

Mobile, on-site training took place under an early CAP grant through the National Park Service in the Great Lakes Region. One trainee under that program stated that she felt her understanding and success in metadata development was significantly enhanced because of the on-site aspect of that training.

Most trainees, when returning to the office after a workshop, are faced with a backlog of work that circumvents “jumping right in” to developing metadata records. For one training in the upper Midwest only four out of 19 trainees worked on metadata development within the first two months after attending a workshop. According to trainers interviewed in this study, this breakdown seems fairly typical. As a result several participants called for follow-up activities such as technical support calls, workshop follow-up assignments, and the availability of online training and support materials through individual agencies’ websites. One NBII trainer distributes cards to each trainee at the end of her workshops that asks, “What are you going to do regarding metadata development when you get back to your office?” She encourages everyone to jot something down even if it’s as simple as reviewing posted metadata records or creating a template.

When successful metadata developers are asked why they are successful while their fellow trainees may not be, they often cite that they’ve been in the position of having to locate data and hence, understand the value of good metadata records. One trainer stated that his first exercise in workshops is to have the trainees search for datasets through a FGDC clearinghouse. This, he stated, helped students grasp the importance of good metadata records.

Two study participants reported that their organization includes metadata training as a part of a one-week data management training workshop. Indeed, metadata is intrinsically tied to data management. In fact, trainers have frequently heard comments such as, “Now I know how to organize my data,” from metadata workshop participants.

Key findings:

1. FGDC’s metadata training program is well-received.
2. Trainees, in general, feel a need to expand the training in scope and duration.
3. On-site (mobile) training has the added benefit of “broadcasting” the value and requirements to the trainees’ co-workers, including their supervisors.
4. Follow-up activities or monitoring after metadata training workshops is reported to improve metadata learning and success rates.
5. Those who have been in the position of searching for and accessing data via FGDC metadata records are more likely to be successful metadata developers.

4.3.2 Higher Education

Interviews with participants employed by colleges and universities showed that very little is being done in the way of metadata training in institutions of higher education. Even within GIS curriculums, metadata is seldom included in coursework. The academic participants of this study warmly welcomed the idea of incorporating metadata in undergraduate or graduate classes.

To facilitate metadata training in academic institutions a few preliminary steps must be initiated. First, a curriculum and course materials must be developed and made widely available. As one participant noted, “Try screening introductory GIS textbooks and see how many include “metadata”.” The fact of the matter is that very few do. Secondly, as in the natural resource management community, supervisors (in this case Assistant Deans) must be convinced of the value of including metadata training in an academic

curriculum. “There’s a problem that most Associate Deans don’t know much about metadata so don’t encourage its’ incorporation into the curriculum,” stated one professor.

Including metadata training at the university level seems to be a logical path to pursue. This author, in conversation with metadata trainers, notes that current training efforts while valuable, do not reach the vast majority of potential metadata developers. Imagine how many scientists, researchers, managers, GIS specialists and other data specialists are in need of training within the numerous local, state, regional and national land management agencies. Divide that number by 20 students per workshop! And, what percent of those 20 will actually produce metadata?

Key findings:

1. Higher education institutions do not incorporate metadata development into current resource or data management curriculums.
2. Higher education institutions are willing to explore how metadata can be incorporated into resource or data management curriculums.
3. Some higher education administrators may need to be convinced of the value of including metadata development in course work.
4. No standard metadata curriculum or course materials are readily available to college and university instructors.

4.4 Metadata Authoring and Presentation Tools

4.4.1 Authoring Tools

Currently, most metadata developers can be considered “occasional” developers. They are generally the data owners and only work on a few research projects per year that involve data collection and management. As time goes on and more data and information management positions are created it may be that this group of biologists, graduate student researchers, GIS specialists, volunteers, etc. need not concern themselves with metadata development.

While current metadata authoring tools are a vast improvement over earlier versions, most study participants reported great difficulty in using them and used terms like, “frustrating”, “inadequate” and “cumbersome.” Those proficient in metadata development had less concerns, but most metadata developers are not proficient. One participant asked, “Why is it that training is required to successfully use SMMS and ArcCatalog?” Several organizations reported developing their own authoring tool in an effort to simplify the metadata creation process.

Participants had recommendations for the next generation of metadata authoring tools. Those recommendations include:

- Framework tools that streamline the process to reduce time and increase quality and frequency.
- An interactive, interview-based interface that includes pop-up windows with suggestions, tips and “do you mean...?” prompts.
- Tools with better integration of metadata with data.
- Open source availability
- Pull-down lists and ability to copy “chunks” of metadata.
- Automated entry of some fields
- Better help linkages and embedded definitions (e.g. press F1 or right click for element definition)

Key Findings:

1. Currently available metadata authoring tools are complex for most biologists, researchers and others being asked to develop metadata records.
2. Beginning and novice metadata developers are seeking easy-to-use, open source tools that require little or no training.

4.4.2 Presentation Tools

Several participants pointed to the idea that the presentation of metadata through clearinghouses is not particularly attractive or user-friendly (easily understandable and usable). They felt the inferior presentation of the end-product, metadata, had a limiting influence on its development. As one data manager stated, “Good pictures will have you pick up the can of beans and read the nutritional label.”

Another stated, “It should not be that users can use metadata through a formal metadata record. Pieces of the metadata record should be extracted into more friendly and usable formats such as a browse list, a catalog of some kind, or even a web page.” A few commented that success is reached when the users don’t know that they’re looking at metadata.

Some offered preferred attributes of metadata presentation. Having the ability to view partial records (thumbnails) before downloading the entire record was noted as was the idea that metadata records could be tied to datasets and through a right click the relevant metadata (not the entire record) would be brought up for viewing. The underlying issue is that change has occurred in how we search for and access information. No longer is personal contact as frequent. Applications need to address the users’ information pathway from defining their question to using the results. Metadata presentation and viewing is a critical element in this pathway. One of the objectives is getting people to understand what they are looking for and a useful presentation of metadata can assist with that which, in turn, will bring higher value to metadata and a greater appreciation of the need for metadata by the developers. One accomplished metadata developer remarked, “Metadata needs to be educational and playful to be used.”

Participants noted a few clearinghouses as having a “friendly looking” and an easier to use metadata presentation; Those mentioned included; Geodata.gov (<http://www.geodata.gov/gos>), the Minnesota Department of Natural Resources (http://maps.dnr.state.mn.us/deli/data_catalog.html), and FAO’s GeoNetwork (<http://www.fao.org/geonetwork/srv/en/main.search>)

Key Findings:

1. Metadata developers feel that improving the visual appeal and usability of the end product will improve the willingness and frequency of developing metadata records.

4.5 The FGDC Standard

With regards to the FGDC standard, four distinct issues came up; definitions for the elements, the first impression of the standard, vocabularies for keywords, and the transition to ISO.

4.5.1 Element Definitions

The terms “vague”, “obtuse” and “jargon” were cited by participants when describing the definitions for the standard’s elements. This was especially true for attribute accuracy and logical consistency elements as those are not commonly addressed in everyday work. Confusion over what exactly is required occurs in other elements as well. For example, trained metadata developers commented that they were not sure what the term “published” meant and that influenced their metadata input. During an observation of metadata entry, two highly experienced developers entered significantly different metadata for the same dataset, both working from the same definitions. Participants called for improved element definitions throughout the standard. The idea of using real world examples within the definitions was mentioned on a few occasions. Experienced metadata developers commented that they frequently have to look up definitions.

4.5.2 First Impression

“Onerous” was used more than once to identify how participants felt when being introduced to the standard. “Imagine a 200-page book is required.” That feeling however, seemed to fade – or at least be accepted – as their experience with FGDC metadata progressed. While it wasn’t reported as such, the fact that the standard is divided into mandatory and optional elements is likely a factor in alleviating some early anxiety. There were requests that the standard be organized to show what elements are necessary for certain user tasks. For instance, would it be possible to organize schemas each comprised of elements necessary to meet the users needs for locating datasets, to retrieving them, evaluating them for relevance or quality, etc? In other words, organize the elements based on the users’ requirements, not the FGDC sections.

4.5.3 Controlled Vocabulary

Most metadata developers have reported to rely on their own knowledge of a subject matter to identify keywords. The problem occurs when users search clearinghouses utilizing their own keywords, or when another metadata developer enters a different term for the same object. Controlled vocabularies are most often used with taxonomic nomenclature, resource types, geographic names, and subject terms (keywords).¹

With a controlled vocabulary, multiple metadata developers and metadata users will more likely use the same term to describe and locate the same dataset or document. For example, we know that Douglas fir = Douglas-fir = *Pseudotsuga menziesii*. A recent search on a USDA Forest Service publications site presented varying results even though we’re looking for the same information.

¹ An Oxford Plant Sciences Library study (unpublished, 2004) found that when comparing 10 forestry datasets utilizing a combined 1,400 keywords, 90% of those terms were unique.

Table 2: Results of search made February 18, 2005 on USDA Forest Service Treesearch site using “keywords” field. <http://treesearch.fs.fed.us/>

Search Term	Number of citations
“Douglas fir”	19
“Douglas-fir”	1
“Pseudotsuga menziesii”	82

4.5.4 Transition to ISO

Regarding the transition to ISO, participants were either nervous about what may happen to their current FGDC records or were delaying developing FGDC records and were in need of clear information on the transition and its impacts on their FGDC records and metadata work. They also expressed an interest in understanding why ISO is an improvement over FGDC.

Key Findings:

1. Definitions for FGDC elements are vague for even experienced metadata developers.
2. FGDC standard appears complex and intimidating on first introduction.
3. Some requested that the standard be organized by users’ requirements.
4. The lack of readily available controlled vocabularies contributes to mis-cataloguing of datasets.
5. There is little information available regarding the planned transition to the ISO metadata, causing some anxiety.

5. Case Studies: Successful Metadata Development

Three case studies are presented. Each case represents a different level of resources available for data management activities, including metadata development.

5.1 Florida Fish and Wildlife Research Institute

The Florida Fish and Wildlife Research Institute (FWRI) is a regional source of information and expertise relating to the marine and wildlife resources of Florida.

The FWRI mission statement reads: “Through effective research and technical knowledge, we provide timely information and guidance to protect, conserve, and manage Florida’s marine and coastal resources.”

At FWRI, scientific research and information synthesis are accomplished through the cooperative efforts of four core groups, including Information Science & Management (IS&M). IS&M employs 29 full-time employees (plus numerous graduate assistants, contractors, etc.) and has a total operational budget of \$3.3 million.

Since 1999 FWRI has viewed spatial metadata as an information asset that must be managed. Important elements that allow FWRI to effectively manage and utilize metadata include:

1. Selection of FWRI-standard metadata application software (SMMS) and database software (Oracle),
2. Identification of both a *de jure* (FGDC) metadata content standard and *de facto* (FWRI-specific) metadata quality standard,
3. Development of written training materials and procedures
4. Delineation of a metadata “coordinator/editor” role to assure metadata quality, consistency and currentness, and
5. Adoption of intranet and extranet strategies for sharing metadata.

FWRI also developed a Data Access Manual with a purpose to “create a single approach to maintaining, storing and accessing digital data collected and obtained by FWRI scientists and staff.” The manual outlines the FWRI “Project Life Cycle” of seven steps from project start to finish that articulates flow, time and staff involvement. Nearly half of the manual is devoted to metadata procedures and standards specifying workflow, required metadata fields, and default content values for many elements.

Finally, FWRI utilizes standard keyword thesauri for theme, place and taxonomy.

Florida Fish and Wildlife Research Institute Website: <http://www.floridamarine.org/>

5.2 USGS Forest and Rangeland Ecosystem Science Center

The USGS Forest and Rangeland Ecosystem Science Center (FRESC) has created a successful metadata development process thanks primarily to strong management support for standardized data set documentation. In 1998 the Center Director established a center-wide policy stipulating that data sets for studies conducted or supported by FRESC would comply with Executive Order 12906 and USGS BRD Policy #8. Together these policies mandate that geographic and tabular data sets be documented in compliance with the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata and the FGDC Biological Data Profile, respectively. Subsequent FRESC Center Directors have reiterated this requirement. FRESC management also facilitates metadata development through the support of metadata coordinator duties as part of the job description of the GIS Specialist. In addition, metadata workshops are presented to FRESC researchers each year.

When new studies are proposed at FRESC, among the items on an administrative file checklist is a documentation review. The metadata coordinator consults with the principal investigator to verify that metadata will be developed for data sets collected as part of the study. Advice and assistance in metadata development are available at any time during the life of a study. Scientists have begun to include metadata in their proposals as research products. When final reports are circulated for review, the metadata coordinator verifies with the principal investigator that metadata have been completed. If the documentation has not been completed the coordinator works closely with the scientist to create the metadata.

A metadata guide (see [Annex A](#)) is provided to the scientist to help organize the information that is needed for compliant documentation. The scientist returns responses to the questions in the guide to the coordinator, who creates the preliminary metadata record. The record is reviewed by the scientist and the record is edited with changes or additions. When the scientist and metadata coordinator agree that all edits have been completed, a final metadata record is posted to the NBII Metadata Clearinghouse.

A successful metadata program depends on three critical factors:

1. strong management support and commitment by researchers,
2. training, and
3. collaboration between the metadata coordinator and scientists.

FRESC Website: <http://fresc.usgs.gov/>

5.3 St. Croix National Scenic Waterway

The National Park Service's St. Croix National Scenic Riverway is the boundary between northern Wisconsin and Minnesota. The Riverway includes 154 miles of the St. Croix River from Gordon Dam, Wisconsin to its confluence with the Mississippi River, and the entire 100-mile-long Namekagon River. A complex mosaic of ownership includes the National Park Service, U.S. Forest Service, St. Croix Chippewa Indians of Wisconsin, the states of Minnesota and Wisconsin, ten counties, and multiple communities and private landowners. At least 238 species of birds visit the area with at least 151 breeding within the park boundary. Insects, amphibians and reptiles thrive in the many niches of the St. Croix and Namekagon rivers. Ninety-four fish species have been identified from the park. Being close to the Twin Cities, the St. Croix River is a popular recreation destination for boating, fishing, bird watching and camping.

The St. Croix National Scenic Riverway's metadata development principally relies on the resident GIS specialist. No other personnel are assigned to conduct metadata activities. Field personnel from St. Croix NSR and partner agencies submit datasets to the GIS Specialist for metadata development and archiving. After reviewing the dataset, the specialist generally contacts the data owner with specific questions about the data. The specialist then begins developing the metadata record and may again contact the data owner with questions of clarification. The specialist usually is working on multiple datasets on any given occasion. Further, this specialist sets aside a block of time to focus on developing metadata records.

After the metadata records are developed and reviewed by the data owner they are submitted for posting to the NR-GIS Metadata and Data Store and the NR-GIS Data Server. <http://science.nature.nps.gov/im/inventory/biology/BiodiversityDataStore>

The GIS specialist cites four principal factors that contribute to her success in metadata development:

1. Supervisory support – the Chief of Resource Management is “convinced of the importance of metadata” and facilitates requisite resources including equipment and time towards metadata development, and data management in general.
2. On-site metadata training – In 1999, through a FGDC CAP grant, a metadata trainer visited St. Croix and other Midwest parks to conduct one-on-one metadata training.
3. Time management – The GIS specialist notes greater success if she sets aside blocks of time to work solely on developing metadata records.
4. The NPS Midwest Office has included metadata development as a criterion for evaluating funding proposals.

St. Croix National Scenic Riverway Website: <http://nps.gov/sacn/management/>

6. Recommended Strategies for Successful Metadata Development within Natural Resource Management Units

Based on the findings from this study, the following recommended strategies are offered for use by natural resource management units. These strategies are organized under three elements identified by many of the study participants as essential for successful metadata development;

1. planning and process,
2. management support, and
3. promotion and education.

6.1 Planning and Process

6.1.1 Develop an Information or Data Management Plan

Information or data management plans provide for a common understanding and approach for handling data and information within an organization. These plans frequently include, but are not limited to, the following topics:

- The organization's mission, goals and objectives especially as related to information and data
- Who the information customers are and their needs
- Information management policies (see below)
- Data and metadata standards
- Data management workflows including metadata development procedure
- Agency-specific and recommended subject-specific thesauri
- Access arrangements/levels and security, including protection of intellectual property and adherence with copyright regulations
- Data and metadata backup procedures
- Performance indicators and progress evaluation
- Personnel and responsibilities

Such plans exist under different names including data access manuals, data management plans, resource management procedure plans, etc. and can be used as models or templates to save time and ensure completeness.

Resources:

National Park Service Inventory and Monitoring Program Data Management with links to data management vision, framework, and standards.

<http://science.nature.nps.gov/im/datamgmt.htm>

Georgia Coastal Ecosystem LTER Site Information System Page with links to data submission guidelines, metadata requirements, data access and use agreements and data storage protocols. <http://gce-lter.marsci.uga.edu/lter/research/guide/gce-is.htm>

Luquillo LTER Site "A Handbook to Data Management Policies, Data Archival and Requests" <http://luq.lternet.edu/datamng/imdocs/division.html>

6.1.2 Create policies and directives to enforce metadata development.

While Executive Order 12906 calls for the creation of metadata for geospatial datasets it is up to the data holding organizations or units to develop working policies for metadata development. These policies can be incorporated into the data management plan. They

do not need to be exhaustive – frequently less than a page is sufficient. They can include requirements for how, when and where datasets and metadata are submitted and archived, data use rights, acknowledgements, etc.

Resources:

Palmer LTER Site Data Management Policy including data submission, access and use.
<http://pal.lternet.edu/lter/dm/datapolicy/>

Executive Order 12906 dictating the requirement for metadata for geospatial datasets.
<http://www.fgdc.gov/publications/documents/geninfo/execord.html>

6.1.3 Include data and information interests within the organization’s mission statement, goals and objectives.

Most resource organizations do not formally recognize data and information as one of their principal assets. This recognition should reside prominently in the text of natural resource organization’s goals and objectives. The following examples illustrate the inclusion of information within an organization’s mission.

“The objectives for our data and information management program are:

1. The data must be persistent, or long-lived.
2. The data must be easy to locate and readily accessible.
3. The data must be of a quality and in a form that is usable, credible and promotes knowledge to a variety of users within and outside our organization.
4. The program must accommodate resource managers’ requirements for accountability.
5. Good data stewardship should be inherent in all resource management activities.”

Lake Mead National Recreation Area Data and Information Management Procedure Plan

“To ensure that all research data collected, synthesized, utilized by or for the organization are scientifically sound, of known quality, and thoroughly documented.”
USFS Pacific Northwest Research Station Information Management Plan

“To improve delivery of information and quality of service, to improve accessibility and usability or resources via a program of training and awareness, and to encourage availability of new, quality networked resources of relevance to the (forestry community).” *Desire Gateway (UK)*

“The NIS identifies the following goals as vital to fulfilling its mission:

- 1) Facilitate management of primary data and information.
- 2) Allow and enhance discovery of information.
- 3) Allow and enhance access to information.
- 4) Facilitate synthesis and integration of information.
- 5) Promote collaboration and community-building.”

Long-Term Ecological Research Network Information System (Draft July 2004)

6.1.4 Re-define the organizational structure and personnel to facilitate metadata.

Managing information is generally considered an “add-on” function to be handled by the data owners. As a result many scientists are thrust into an information management

role. Data and information are intellectual capital and should be treated accordingly. Increasingly, natural resource organizations are shifting the responsibilities of information management to a level of support similar to that of financial management. Ensure that the organizational chart reflects the value of data. Only when personnel see this emphasis will they begin to respect the value of data – and metadata.

To accomplish this structure an information specialist or data manager needs to be assigned or hired and given a position of recognized responsibility. Numerous organizations maintain a full complement of biologists and GIS specialists, but not all maintain someone to manage the information produced by these scientists. Study participants that worked for organizations that regularly created metadata records unanimously stated that having an individual responsible for information management was a priority. It was reported that these individuals conduct a variety of metadata responsibilities including:

- Arrange and conduct metadata training
- Create and/or assist with the development of metadata records
- Review and edit metadata records for quality and conformity with organizations schema
- Submit datasets and metadata to appropriate archives/clearinghouses
- Develop metadata development forms and project cycle process steps with responsibilities
- Identify and maintain metadata schema, authoring tools and keyword lists
- Promote metadata development to staff
- Assist with access to metadata records

Resources:

Oregon Bureau of Land Management - Data stewardship role to support data management. <http://www.blm.gov/nhp/efoia/or/fy2003/ib/ib-or-2003-037Attach1.htm>

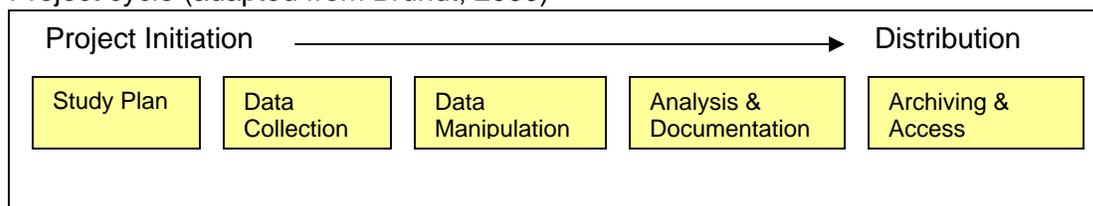
NPS Inventory and Monitoring Data Management site with link to data managers responsibilities. <http://science.nature.nps.gov/im/datamgmt.htm>

6.1.5 Establish a metadata development process that corresponds to the project life cycle.

Ensure that metadata development begins before data collection activities and is conducted throughout the life of the project. Also ensure that metadata development responsibilities are shared among personnel corresponding to the data management process.

Evidence strongly supports that metadata development is facilitated if it is incorporated into the design and implementation phases of a project. This means that the initial development of metadata needs to occur in the study plan phase of a project.

Project cycle (adapted from Brundt, 2000)



As the project progresses periodic updating of the metadata record should be made. If changes such as authors, collection sites or protocol, software used for analysis, or any other aspect of the project are made, those changes must be reflected in the metadata record. In the end, the metadata record should be an integral (“attached to”) element of the final document or dataset.

Resources:

Wayne, L. “Institutionalize Metadata *Before it Institutionalizes You*” URISA 2001.
http://www.fgdc.gov/metadata/education/InstitutionalizeMeta_URISA2001.pdf

Brundt, J., 2000. “Data management principles, implementation and administration.” In: *Ecological Data: Design, Management and Processing*. Edited by W.K. Michener and J.W. Brundt, p. 25-47. Blackwell Science, Malden, MA.

6.1.6 Adopt existing recognized controlled vocabularies.

Information experts have found that better consistency is obtained over time and space if standards are used for controlling the terms in an information repository. Semantic control is achieved by identifying existing or developing new vocabularies. These vocabularies can be in the form of keyword lists, thesauri, gazetteers, glossaries, or alpha-numeric lists. Numerous controlled vocabularies exist for use with specific disciplines, e.g. soils, forest hydrology, taxonomy, fire. Once identified or developed, both information contributors and users can be assured that they are referring to the same species, locations, and concepts.

Whatever controlled vocabulary, or combination thereof, is chosen the selection process can be facilitated based on the following criteria:

1. Is this CV used? How often? If not currently used, when planned?
2. What categories of users use the CV? (e.g. researcher, public, etc.)
3. Is the CV well-matched to information your organization is concerned with?
4. Is the CV documented? (Could include bibliography of contributing sources, how CV terms are developed, e.g. committee)
5. Is the CV “mapped” with any other CV? (Sometimes simple associations, e.g. using terms from another accepted glossary.)
6. Does CV exist in machine-processable form? (Delimited text documents, etc.)
7. Who has authority of the CV and how often do they update it?
8. What is the level of scientific detail and validity of the CV?
9. Is the CV a glossary, hierarchical thesaurus, other?
10. Is the CV appropriate to use on a scientific basis?

Resources:

Dublin Core Controlled Vocabulary Tutorial
http://dublincore.org/resources/training/dc-2004/english/DC-2004_Tutorial_4_en.pdf

USGS Keyword Thesauri and Attribute Label Definition Sources (bottom of page)
<http://www.fgdc.gov/metadata/links/metalinks.html>

6.2 Management Support

6.2.1 Educating managers about metadata

Educating managers on the value and process of metadata development is a critical component of a metadata management program. Simple steps can be taken to facilitate support from mid- and upper-level managers that will result in improved metadata resources and records throughout the organization. One method would be to conduct metadata orientation seminars. Such seminars would include the following topics:

- What is metadata
- Why is it needed
- What are the benefits
 - For data users
 - For data providers
- The process of implementing metadata development
- Keys to metadata success
- Metadata resources
- Training opportunities

Many managers want to be assured that the effort and cost put into metadata development will result in significant benefits. Measuring the cost of developing datasets, or in other words the cost of losing datasets, is not a simple task, but one that may be well worth the effort.

Resources:

FGDC Metadata Brochure for Managers

<http://www.fgdc.gov/publications/documents/metadata/metabroc.html>

NOAA IT Training for Managers

http://www.csc.noaa.gov/training/info_tech.html

6.3 Promotion and education

6.3.1 Capitalize on FGDC metadata training opportunities

Attendance at FGDC hands-on metadata workshops is encouraged. It is further encouraged that more than one individual from a management unit should attend. Just as it is important to have metadata development be a multi-person effort, training should include more than one individual per management unit.

If possible, one individual within a management jurisdiction should strive to become a metadata trainer. Those organizations that reported on-site trainers achieved significantly greater success in metadata development. FGDC offers periodic “Train the Trainer” programs.

Resources:

FGDC Tutorials, trainer list and training calendar:

<http://www.fgdc.gov/metadata/education.html>

6.3.2 Develop agency or organization metadata support websites

Several organizations have taken it upon themselves to develop online resources to support personnel engaged in metadata development. These resources include internal metadata and data management discussion forums, FAQ's, links to tools, element definitions, example metadata records, and other pertinent information. Not only does this help agency personnel develop metadata, it also raises the awareness of metadata within the organization.

References:

Marine Metadata Interoperability website: <http://marinemetadata.org/>

Center for International Earth Science Information Network (Columbia University):
<http://sedac.ciesin.columbia.edu/metadata/>

Minnesota Land Management Information Center metadata resource page:
http://www.lmic.state.mn.us/chouse/meta_help.html

6.3.3 Promoting metadata development with recognition and incentives

While no examples of a reward or incentive program were given, study participants expressed a keen interest in the idea of receiving recognition from their supervisors for good metadata development. Recognition could be as simple as a certificate given at a staff meeting or staff lunch to honor the good metadata developer. Metadata development is, after all, a somewhat tedious job and as with all tedious jobs appreciation for work accomplished is warmly received.

References:

Wilshire Award for Best Practices in Meta Data Management:
<http://www.wilshireconferences.com/award/index.htm>

6.4 Evaluating and Prioritizing Recommendations – An Example

Ten recommended strategies for enabling metadata development have been made. It is recognized that priorities must be set when embarking on new recommendations such as those in this report. To be effective these recommendations must, in some way, contribute to the organization's data management objectives. Due to differences in missions (research, resource utilization, preservation, tourism, etc.) natural resource organizations maintain different data management objectives. Table 3 shows an example chart for subjectively evaluating and prioritizing the contribution that these 10 recommendations have on Lake Mead NRA's stated data management objectives (see [6.1.3](#)).

Table 3. Relative contribution of recommendations on DM objectives of Lake Mead NRA (example only).

Recommendation	Persistent	Accessible	Credible	Accountable	Stewardship
6.1.1 Information Planning	✓✓	✓✓	✓✓✓	✓✓	✓✓✓
6.1.2 Policies & Directives	✓✓	✓✓	✓✓	✓✓	✓✓✓
6.1.3 IM in Mission Statement	✓✓	✓✓	✓✓	✓	✓✓✓
6.1.4 Organizational Structure & Personnel	✓✓	✓✓✓	✓✓	✓✓	✓✓✓
6.1.5 Metadata Development Process	✓	✓	✓✓✓	✓✓	✓✓✓
6.1.6 Adopt CV's	✓✓	✓✓✓	✓✓	✓	✓✓✓
6.2 Management Support	✓✓	✓✓	✓✓✓	✓✓✓	✓✓✓
6.3.1 FGDC Metadata Training	✓✓	✓✓	✓✓	✓✓	✓✓✓
6.3.2 Organizational metadata support websites	✓✓	✓✓	✓	✓	✓✓✓
6.3.3 Recognition & Incentives	✓✓	✓✓	✓✓✓	✓✓	✓✓✓

✓ = low ✓✓ = medium ✓✓✓ = high

7. Recommendations to FGDC

Study participants identified several activities that FGDC could carry out to assist them in their metadata development efforts. There was an overwhelming appreciation for what FGDC has accomplished with regards to developing the standard, promoting metadata and metadata tools, and providing training. Most felt that FGDC's role should focus on education and promotion of metadata and shift away from a standards organization.

7.1 Educational

It was frequently stated that education needs to be at the center of FGDC's efforts. In reflecting the common opinion, one participant stated, "Education, education, education," when asked what she would recommend to FGDC. Several mentioned that it isn't technology that is the limiting factor to metadata development, but instead it is educating the broad community of scientists, educators, managers, and other stakeholders of data and information management.

Participants felt that the CAP program is one of FGDC's "shining stars" and most expressed an opinion that the program should somehow be expanded to include more recipients and/or categories. It was stated on numerous occasions that CAP-funded metadata training workshops have been well-attended and there is a growing demand for more training.

In a nutshell, participants noted several ways that the CAP program could be expanded. Their ideas include:

1. Increase the number of hands-on metadata training workshops
2. Facilitate an on-site mobile metadata training (or traveling trainers) program
3. Lengthen the time of metadata training workshops
4. Add funding for follow-up activities to metadata training workshops
5. Add a category for education and promotion activities. Ideas mentioned include:
 - a. Developing a metadata business case
 - b. Compiling and posting metadata FAQ's
 - c. Conducting annual metadata best practices recognition awards
 - d. Conducting "Metadata for Managers" workshops
 - e. Developing curriculums and materials for including metadata training in college and university programs
6. Establish an online metadata educational resource and training center

7.2 Outreach

Besides requesting additional and broader educational programs from FGDC, participants noted that there is a real need for promoting metadata. It was stated that outreach needs to be directed at all levels, from Washington D.C. agency administrators to field level biologists. In particular, mid-level managers were frequently identified as potential recipients of metadata education and outreach. These individuals allocate personnel, financial and equipment resources for day to day activities that include data management. If they are not convinced of its' importance, advancement in metadata development will continue to be a struggle.

Stated promotional activities that were recommended to FGDC include:

1. A proactive approach including newsletters and/or e-mail announcements to stakeholders, especially to state lists, about FGDC clearinghouses and other ongoing activities.
2. Frequent contact with resource management agency headquarters personnel
3. Seeing the FGDC community as a metadata community with a goal of getting data and information to users, e.g. less focus on the FGDC standard
4. Improving the FGDC website's "look and feel" (no specifics given)
5. Promoting the ISO standard
6. Publishing an engaging metadata poster for resource management units – could show metadata lifecycle

It was felt that most of the potential audience of FGDC services was not even aware of FGDC and the services they have to offer. "Currently, only those that know about metadata will access the FGDC website," was a revealing comment. Especially when one considers all the individuals at parks, ranger districts, research stations and other resource units that literally don't know what metadata is.

One participant mentioned that the institute she works for have annual campaigns to get metadata promoted and developed. She recommended a similar campaign approach on a national scale. Their campaigns focused on a particular species or issue, such as invasive species or water quality, rather than focusing on the broader and less pressing (to the data owners) issue of metadata development. An example may be a "Cataloguing Western Birds" or a "Preserving the Nation's Native American Cultural Legacy" campaign. In some sense, such campaigns would simply be placing metadata development in the context of more significant and appealing issues – a part of a more marketable package.

Some participants noted that recognition is one of their greatest rewards within their career of biological sciences. The lack of recognition for metadata development plays a role in their willingness to develop metadata. The idea of being the next John Muir had an appeal that led them to study science and carries them in their career. One such recognition model may be the Wilshire Metadata Award given annually to private companies for outstanding metadata management practices. Previous winners include Allstate Insurance, Intel, and Bell South. See <http://www.wilshireconferences.com/award/index.htm>

7.3 Tools

Interviewees that develop a significant number of metadata records were quite comfortable with current metadata authoring tools, including SMMS, ArcCatalog, Corpsmet and tkme. Other participants, however, consider themselves occasional metadata developers. These individuals are generally scientists and the "owners" of the dataset to be catalogued. These occasional metadata developers reported that current authoring tools are too complex – even after attending a metadata training workshop.

Ideas for how FGDC could assist in tool development were expressed. Most felt that an easy-to-use (e.g. training not required) interactive wizard-style program should be available. They stated that an interview-based system such as those used by popular tax preparation programs would be appealing. There was differing views on whether FGDC should facilitate privately held or open source authoring tools.

7.4 The Metadata Development Process

Participants pointed out on several occasions that there were some components of metadata development process that FGDC could assist with. These components are activities that metadata developers either did not have the time or expertise to conduct.

1. Sample templates that could be modified and utilized in their own metadata records.
2. A generic business plan to enable them to communicate to their managers the benefit of metadata development.
3. A generic metadata management plan that outlines the process for developing metadata as well as follow-up maintenance requirements.
4. Sample data management and metadata development policies that could be modified and adopted for use within their organization.

7.5 The Standard

Participants simply noted in several instances that their problem with the standard is not as much its complexity as it is the definitions for most elements. They would like FGDC to update the definitions and where possible include “real world” example entries. Other methods for assisting developers with definitions were noted including maintaining a FAQ site for definitions and a listserv to answer definition questions.

Some felt that the FGDC schema should be re-organized to more closely reflect the users’ needs and applications.

7.6 Other Recommendations

Other recommendations were made by participants that do not fit in the above categories include re-invigorating the FGDC committee process and holding a meeting that would bring together experts in the metadata domain to discuss how to proceed in metadata education and implementation.

With regards to the FGDC committee status, a few participants commented that the current committees are less effective than intended, but felt that the process is important to continue. A committee for standards and one for evaluation were mentioned as candidate committees.

Finally, it was suggested by two participants that a FGDC-sponsored “Metadata Summit” be held. It was envisioned that such a meeting would touch on topics that are addressed in this study. It was also envisioned that the meeting would be a roundtable discussion of ideas and not include presentations and be limited to 10-15 metadata practitioners and experts.

Annex A: FRESA Electronic Metadata Interview

(Enable MS Word “comments” or “markup” feature to view FGDC element applications)

1. Have you already prepared metadata for this data set?
 - a. If yes, please send a copy of the documentation or a reference to where it can be found and skip to item **17**.
2. What is the title of the data set?
3. Who is the originator(s)/ owner of the data set? (Include address and telephone number)
 - a. If someone else should answer questions about the data, please list the name, address, and telephone number.
 - b. Are there other organizations or individuals who should get credit for support, funding, or data collection and analysis?
4. Is the data set published or part of a larger publication?
 - a. If so, what is the reference?
5. Include a brief (no more than a few sentences) description of the data set.
6. Why were the data collected in the first place?
 - a. Do you have any advice for potential users of the data set?
7. What is the time period represented by the data set?
8. Was the time period determined primarily through:
 - a. Field visits
 - b. Remote instrumentation (*i.e.* temperature recorders, etc)
 - c. Existing data sources
9. What is the status of the data set you are documenting? – *complete, in progress, planned*
 - a. Will the data set be updated? If so, how frequently?
10. Where were the data collected? Include description and coordinates, if known.
11. List some keywords to help search for this data set.
 - a. Thematic, Place, Temporal, Strata, Taxonomy
 - b. If a controlled vocabulary was used, what is the reference?
12. Data Access
 - a. Are there policy or legal restrictions on who may access the data?
 - b. Are there such restrictions on how the data are used?
13. The FGDC Biological Profile includes the means to document tabular data sets, taxonomy, field methods, and the use of analytical tools or models.
 - a. Was your data set developed using a model or other analytical tool?
 - i. If so, describe how the model works
 - ii. What is the reference?
 - iii. If the model or tool is available include a contact and/or URL
 - b. Does the data set contain biological information? If no, skip to item **14**.
 - i. What species or communities were examined?
 - ii. Did you use a taxonomic authority or field guide for identification?
 1. If so, what is the reference?
 - iii. Briefly summarize your field methods (copy&paste from other documents!)
 1. If you used existing protocols or methods, list the references
14. Is your data set archived in a databank or data catalog? If yes, please include a reference to the documentation and skip to item **15**. If No:
 - a. What tests did you conduct to be certain your attributes actually represent field observations? Date of tests?

- b. What measures did you take to make certain that your data set was as nearly correct as possible?
 - c. Were there any things that you excluded from your data collection? *i.e.* stems less than a certain diameter or streams without surface flow
 - d. What is the form of your data set? - *spreadsheet, ascii file, gis layer, database, other*
 - e. What is the filename for your data set?
 - i. For each file or table, list the fields in the data set and for each field list:
 - ii. The definition of the field
 - If the data are coded (categorical), what are the definitions?
 - 1. are the codes unordered (nominal)? or
 - 2. are the codes ordered (ordinal – low, med, high)?
 If the data are measured, what are the allowable min and max?
 - 3. are they from a relative scale with no natural zero (10 ° C)? or
 - 4. are they from an absolute scale (10.5 meters)
 Are the data date or time values? List the format *i.e.* mm/dd/yyyy
 - iii. If the codes come from a published code set, list the reference
 - iv. Otherwise, the domain is unrepresentable. Include a brief description of what is in the field
15. Is this a GIS data set? If no, skip to item **16**.
- a. Send a copy of the ArcInfo export file, an ArcView shapefile, or an ArcCatalog exported metadata file (txt or xml).
 - i. Include projection parameters, if necessary.
 - b. List any source data sets you used. For each source list:
 - i. Source name, originator and publication date
 - ii. Source time period and scale
 - iii. Source presentation form and media type
 - iv. Contribution of source to your analysis
 - c. List the processing steps you used to create your data set, including the approximate date of processing.
16. Is the data set available for distribution? If no, go to **17**
- a. What are your distribution instructions?
17. You are done. Send a text file containing the responses to this interview to your metadata coordinator.