

# Metadata Standard Operating Procedure

## Revision Log

| Revision | Description of Change                                 | Author     | Effective Date |
|----------|---|------------|----------------|
| 1        | Original draft  | G. Sanders | 10/01/2005     |
| 1.1      | Updated info re: metadata tools                       | G. Sanders | 09/14/2006     |
| 2.0      | Redrafted most of the documentation                   | G. Sanders | 03/16/2007     |
| 2.1      | Added Metadata Questionnaire and additional guidance. | G. Sanders | 01/23/2009     |

## Purpose

The purpose of this standard operating procedure is to familiarize NPS staff and cooperators with the requirements that the NCRN I&M Program has for dataset documentation. This document is also meant to provide a minimal amount of guidance for metadata.

## Scope and Applicability

The NCRN requires that all datasets, both spatial and non-spatial, be accompanied by metadata that meets the Federal Geographic Data Committee (FGDC) metadata standards. Accordingly, any dataset (spatial or non-spatial) produced by or for the NCRN must be accompanied by FGDC compliant metadata.

This document is not intended to serve as step-by-step instructions for metadata generation, but should serve as an overview of metadata requirements and provide sources that can guide the reader through the process of generating FGDC compliant metadata.

## Acronyms and Definitions

|                                |  |
|--------------------------------|--|
| <i>Arc Catalog</i>             | Module in ESRI's ArcGIS software within which metadata for spatial data sets can be created.   |
| <i>Biological Data Profile</i> | Set of definitions for the documentation of biological data through the creation of extended elements to FGDC Content Standard.                          |
| <i>CSDGM</i>                   | Content Standard for Digital Geospatial Metadata.  |
| <i>ESRI</i>                    | Environmental Systems Research Institute – GIS software company  |
| <i>FGDC</i>                    | Federal Geographic Data Committee – Interagency committee that promotes the coordinated development, use, sharing, and dissemination of geographic data. |
| <i>Metadata</i>                | Data about the content, quality, condition and other characteristics of a data set, documented in a standardized format.                                 |
| <i>NCRN</i>                    | National Capital Region I&M Network  |

## Metadata Standard Operating Procedure

|                       |  |
|-----------------------|--|
| <i>NPS Profile</i>    | The NPS Natural Resources Profile extends the FGDC CSDGM to incorporate NPS-specific elements such as park unit.                             |
| <i>NPS Data Store</i> | Web-based clearinghouse for GIS and tabular datasets. All data sets posted to the Data Store must be accompanied by FGDC compliant metadata. |
| <i>XML</i>            | Extensible Markup Language. A simple and flexible text format that facilitates large-scale electronic publishing and exchange of data.       |

### Reference Documents

Code of Federal Regulations, Executive Order 12906 (1994) Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure. Document available online at <http://www.fgdc.gov/publications/documents/geninfo/execord.html>

OMB Circular No. A-16. Document online:  
[http://www.whitehouse.gov/omb/circulars/a016/a016\\_rev.html#5](http://www.whitehouse.gov/omb/circulars/a016/a016_rev.html#5)

Federal Geographic Data Committee (1998) Content standard for digital geospatial metadata. Document available online at: <http://www.fgdc.gov/metadata/csdgm/>.

National Spatial Data Infrastructure (2004) Metadata Factsheet. Document available online at <http://www.fgdc.gov/publications/documents/metadata/metafact.pdf>.

NPS Metadata Tools and Editor: Creating Non-Geospatial Metadata for the NPS Data Store. Document online: <http://science.nature.nps.gov/nrdata/docs/metahelp/NPSDataStoreCreatingNGSMetadata.pdf>

NPS Metadata Tools and Editor: Creating Simple Geospatial Metadata. Document online:  
<http://science.nature.nps.gov/nrdata/docs/metahelp/NPSDataStoreCreatingGSMetadata.pdf>

Biological Profile (National Biological Information Infrastructure – NBII) Metadata Guide. Document online:  
<http://science.nature.nps.gov/nrdata/docs/metahelp/BiologicalProfileGuide.pdf>

NPS Data Store: Metadata and Data Uploading Guidance. Document online:  
<http://science.nature.nps.gov/nrdata/docs/metahelp/NPSDataStoreMetadataDataUploadGuidance.pdf>

NPS Data Store: Parsing Metadata with NPS Metadata Tools and Editor. Document online:  
<http://science.nature.nps.gov/nrdata/docs/metahelp/NPSDataStoreMetadataParsingGuidance.pdf>

### Procedures and General Requirements

#### *Overview*

Data documentation is a critical step toward ensuring that data sets are usable for their intended purposes well into the future. This involves the development of metadata, which is defined as structured information about the content, quality, condition, and other characteristics of a data set. Basically speaking, the 'who', 'what', 'when', 'where', 'why' and 'how' of the data set. In addition, metadata include information about data format, collection and analysis methods, access/use constraints, and distribution. Metadata provide the

## Metadata Standard Operating Procedure

means to catalog data sets, within intranet and internet systems, making their associated data sets available to a broad range of potential users. While most frequently developed for geospatial data, metadata describing non-geospatial data sets are also needed. For example, water samples collected daily for an annual report to summarize water quality should be documented with complete protocols and metadata for the database in which the data are stored.

Executive Order 12906, issued in 1994, established the National Spatial Data Infrastructure (NSDI) and, in an effort to strengthen the policies outlined in OMB Circular No. A-16, directed agencies to “document all new geospatial data it collects or produces”. The NPS extends these policies to include all *spatial and non-spatial data sets*.

The general goal of the NPS metadata system (Figure 1) is to catalog all data sets and to produce FGDC-compliant metadata for those data sets that require comprehensive documentation. The NPS Natural Resource, GIS and I&M Programs released the NPS Data Store in 2005 as an FGDC-structured database and data server system that provides a secure web interface and tools to import metadata records from desktop metadata authoring programs.

The NPS Integrated Metadata System Plan for spatial and natural resource data sets incorporates the generation of metadata by Network and Park users (through desktop applications), metadata posting to and retrieval from the NR-GIS Metadata and Data Store (through an online, Oracle-based server), and dissemination of non-sensitive records to the public through the Geospatial One-Stop clearinghouse available at: <http://gos2.geodata.gov/wps/portal/gos> (Figure 1). When completed, the metadata database component of this system will become the master database for NPS spatial and natural resource metadata. For more information, see the NR-GIS Metadata and Data Store website available at: <http://science.nature.nps.gov/nrdata/docs/metaplan.cfm>.

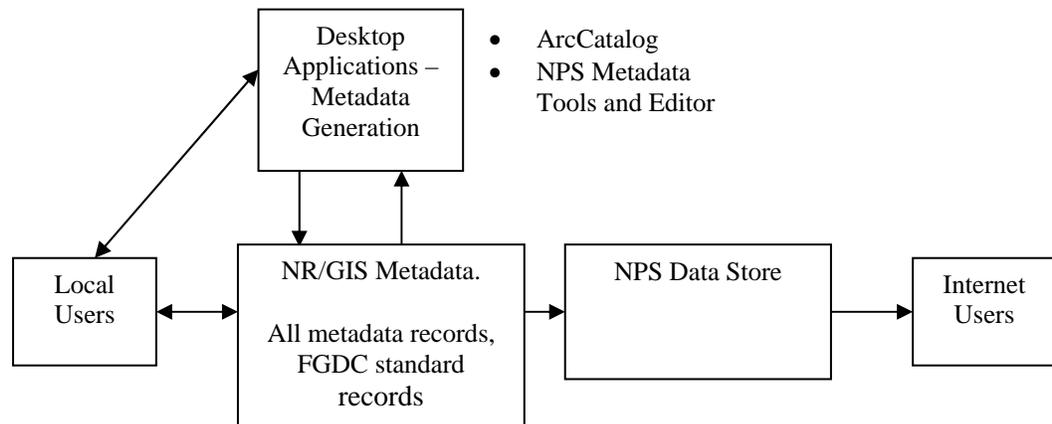


Figure 1. Metadata workflow diagram, adapted from the NPS Integrated Metadata System.

## Metadata Standard Operating Procedure

There is no single, standard software package used for metadata generation within the NPS I&M Program. The two programs that the NCRN strongly recommends are used are: the NPS Metadata Editor and ArcCatalog and they are described below. The functionality of the tools are outlined in Table 1.

- *NPS Metadata Tools & Editor*: The I&M Program has developed a Metadata Editor for editing and managing metadata that can function as a stand alone application or as an extension of ArcCatalog. The application is intended to be the primary editor for metadata that will be uploaded to the NPS Data Store. The NPS Metadata Tools and Editor application integrates with the NPS Data Store information system by producing XML metadata based on the NPS Metadata Profile. This XML metadata file can then be uploaded to the NPS Data Store application.
- *ArcCatalog*: ArcCatalog is part of ERSI's ArcGIS® software system. With ArcCatalog, you can browse, manage, create, and organize your geographic and tabular data. In addition, ArcCatalog comes with support for several popular metadata standards to allow you to create, edit, and view information about your data. There are editors to enter metadata, a storage schema, and property sheets to view the data. ArcCatalog may be used for metadata generation for all geospatial data.

Table 1. Comparison of the three main metadata generation tools.

| Tool/Features              | ArcCatalog       | NPS Tools & Editor      |
|----------------------------|------------------|-------------------------|
| Minimal Metadata           | Y                | Y                       |
| Comprehensive Metadata     | Y                | Y                       |
| Biological Metadata        | N                | Y                       |
| Create metadata templates? | Y                | Y                       |
| FGDC Section 1             | Y                | Y                       |
| FGDC Section 2             | Y                | Y                       |
| FGDC Section 3             | Y                | Y                       |
| FGDC Section 4             | Y                | Y                       |
| FGDC Section 5             | Y                | Y                       |
| FGDC Section 6             | Y                | Y                       |
| FGDC Section 7             | Y                | Y                       |
| Import                     | .sgm, .txt, .xml | .sgm, .txt, .xml        |
| Export                     | .txt, .xml       | .txt, .xml, .sgm, .html |
| Parsing                    | N                | Y                       |

## Metadata Standard Operating Procedure

- *NPS Database Metadata Extractor*: The Metadata Extractor is Microsoft Access Add-in developed by the Natural Resource GIS Program. The tool is designed to harvest content information from MS Access database including information on database entities (tables), attributes (fields each table) and domains. The application allows the user to edit the harvested information (including batch edits) and export it as an FGDC compliant XML file. The XML file generated can be incorporated into an existing metadata template using the Metadata Tools & Editor to provide content for section 5 of the template. <http://science.nature.nps.gov/nrddata/tools/dme.cfm>
- *Metadata Questionnaire*: NCRN staff understands that not everyone is familiar with metadata or the two metadata generation tools previously mentioned. Metadata is still a programmatic requirement and it is very important to have the appropriate staff and/or cooperators involved in the data documentation process. Therefore, the NCRN has adapted a metadata questionnaire (KLMN 2007; NCCN 2006) for project staff to complete which will provide NCRN staff the information needed to update the electronic metadata file. A copy of the questionnaire can be found in Appendix A.

### *Metadata Generation*

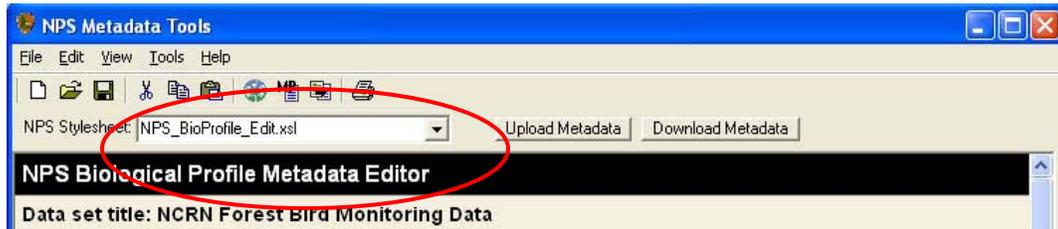
Metadata creation should begin during the planning phase of a project. During this time the Data Manager will create a metadata "template" for the project wherein much of the basic information about the project such as network contacts, associated parks, and keywords are added. The metadata record should continue to be updated during the data acquisition and data delivery phases of a project.

The standard FGDC CSDGM consists of seven sections (1-7). The NPS includes as additional section (0) in the NPS Metadata Profile provided with the Metadata Tools and Editor. Section 0 maintains information about the park units associated with the data set, NPS specific key words and identifies the NPS data steward.

Of the eight sections in the metadata profile, the minimum that the NPS requires are sections (described in more detail below): 0, 1, 6, and 7. Staff and cooperators are encouraged to provide additional and more detailed information in the other sections especially Section 2 where specific information about methodologies can be included.

- Metadata must be created for every GIS data layer as well as all tabular datasets.
- A metadata file (.xml) should be created at the beginning of every project and updated as the project progresses.
- When creating/editing metadata for non-spatial data files using the NPS Metadata Tools and Editor, use the BioProfile Style Sheet (NPS\_BioProfile\_Edit.xsl; see Figure 2 below). This style sheet will allow you to create/edit metadata for Section 2 (Data Quality).

## Metadata Standard Operating Procedure



**Figure 2.** Screen capture from the NPS Metadata Tools & Editor illustrating the proper metadata style sheet to use.

### *Review for Sensitive Data*

All data sets must be reviewed to see if they contain any sensitive information. Sensitive information could include information on any of the following things:

- rare, threatened, endangered, or commercially valuable species
- rare habitats or habitats that are essential to an RTE species
- sensitive resources such as caves or wells
- minerals
- paleontological objects
- archeology resources
- ethnographic sites

The project manager should review all of the data and materials and consult with network staff and members of the field crew to confirm whether any sensitive information is present. Cooperators and contractors conducting any monitoring protocols monitoring should also be asked to provide any input on the possible existence of sensitive information. Finally, natural resource managers and other park staff should be asked to review the protocol data and products for the presence of sensitive information.

The following questions can aid in determining whether information is sensitive:

- Has harm, theft, or destruction occurred to a similar resource on federal, state, or private lands?
- Has harm, theft, or destruction occurred to other types of resources of similar commercial value, cultural importance, rarity, or threatened or endangered status on federal, state, or private lands?
- Is information about locations of the resource in the park specific enough so that the park resource is likely to be found at these locations at predictable times now or in the future? (For example, raptor nests locations).
- Would information about the nature of the park resource, if available in conjunction with other public information, permit determining specific locations of the resource?
- Even if relatively out-dated, is there information that would reveal locations or characteristics of the park resource such that the information could be used to find the park resource as it exists now or is likely to exist in the future?
- Is the NPS unable to protect the park resource if the public knows or can infer its location?

## Metadata Standard Operating Procedure

- Is the information temporarily critical to some park's operations or negotiations (for a land purchase, water rights issue)?

Answering yes to any of the previous questions qualifies the information as sensitive. Any ambiguity in how the data should be classified results in the data being classified as sensitive.

*Metadata Outlined (sections in bold are the minimally required sections)*

- **Section 0: NPS Information:**

This section is only found if using the NPS Metadata Tools and Editor application (either as stand alone or incorporated into ArcCatalog). This section identifies the park units associated with the data as well as specific keyword categories and contact information required by the NPS Data Store. This section will normally be completed by NCRN staff and included in the 'Template' metadata document provided to the cooperators.

- **Section 1: Identification Information:**

This section contains mandatory information such as citation details, data set abstract and purpose statements, relevant time period associated with the data set, contacts and keywords. Other important details to include in this section are the status of the data set and how often the data set is updated. If the data set contains sensitive information there is a field in this section where use constraints can be included that limit the distribution of the data set.

**Section 2: Data Quality Information:**

This section, while not mandatory, it is highly recommended that it be completed and is accessible by using the 'NPS\_BioProfile.xml' style sheet. If the data set contains taxonomic information, such as species lists, that information can be entered in this section. Entering that information can be incredibly tedious especially if the list of species is lengthy. The NPS Metadata Tools and Editor includes a tool for importing this information. Please refer to the [Biological Profile Metadata Guide](#) for step-by-step instructions.

In addition to taxonomic information the Biological Profile also includes fields for inputting information about field and lab methodologies as well as details on analytical tools (e.g. statistical procedures) that were used to produce data products. If the data set was derived from raw data this section should document the original data sources and describe any procedures used to develop the derived data products.

**Section 3: Spatial Data Organization Information:**

This section only applies to spatial data sets and contains information on the mechanism used to represent spatial information in the data set. ArcCatalog automatically harvests this information from spatial data sets.

- **Section 4: Spatial Reference Information:**

This section only applies to spatial data sets and contains a description of the reference frame for, and means of encoding, coordinates in the data set. Examples include the name of and parameters for map projections or grid coordinate systems, horizontal and vertical datums, and the coordinate system resolution. This information is also automatically harvested from spatial data sets by ArcCatalog.

## Metadata Standard Operating Procedure

- Section 5: Entity and Attribute Information:

This section contains information about the content of the data set, including the entity types and their attributes and the domains from which attribute values may be assigned. Examples include the names and definitions of features, attributes, and attribute values. The NPS Metadata Tools and Editor does not include a template for section 5 but it can be incorporated into a metadata file by using the 'Update with XML Template' tool. Metadata for this section can be generated (FGDC compliant XML) from project databases (MS Access) using the Metadata Harvester (described above)

- Section 6: Distribution Information:

This section provides information about obtaining the data set. Examples include contact information for the distributor, available formats, information about how to obtain data sets online, and fees for the data. This section also contains the standard NPS Liability Statement included by default into every NCRN metadata record.

- Section 7: Metadata Reference Information:

This section deals specifically with information pertaining to the metadata record itself. It describes, who created it, when it was created, when it was updated and all of the contact information for the person responsible for the metadata. This section will be completed and distributed with the 'template' metadata file but as cooperators or staff update metadata records, they must indicate the date when the metadata record was updated.

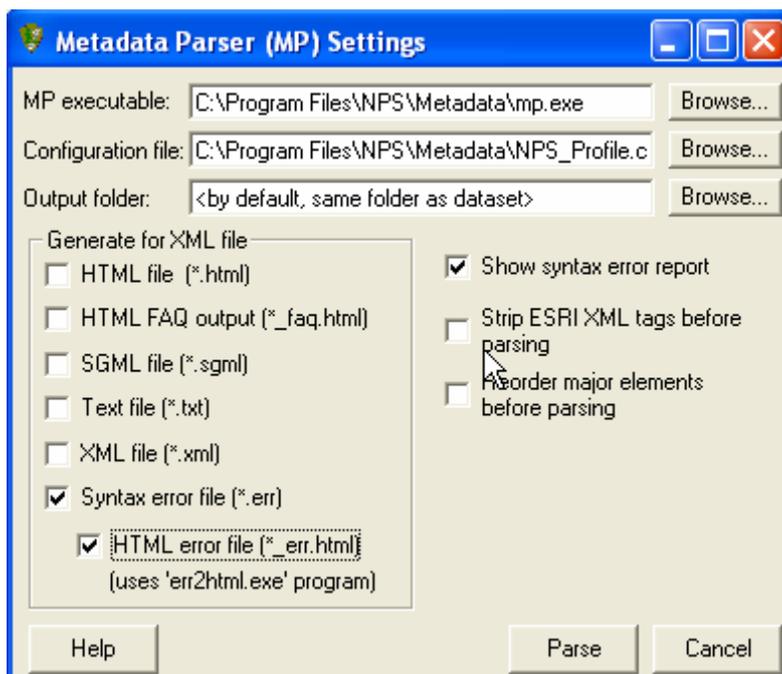
### *Metadata Completion*

When the project is complete, or in the case of long-term monitoring datasets, when data set milestone is reached (e.g. annually), the metadata record should be finalized by reviewing for quality, FGDC- and NPS-compliance, and parsing for errors. The metadata and data set must also be reviewed by cooperators and/or NPS staff to determine if the data set or metadata contain any sensitive information. If the data set does contain sensitive information make sure that the Section 1 of the metadata record addresses this.

### Parsing Metadata:

The metadata file must be parsed using the Meta Parser built into the NPS Metadata Tools and Editor. This will identify any incorrectly placed metadata elements or omissions in the metadata record. To parse a metadata record select the 'Parse with MP' tool from the Tools menu. The Metadata Parser dialog box will open and be sure to select the appropriate settings (Figure 2).

## Metadata Standard Operating Procedure



### **NOTE:**

It is always a good idea to save a copy of your metadata record prior to parsing it with the Metadata Parser.

Figure 2. Metadata Parser settings dialog box.

The Metadata Parser will produce a series of error reports. These errors should be addressed as best as possible. In some cases some errors will be unavoidable as is the case when a non-geospatial data set has errors due to missing spatial domain information. In such cases, errors should be ignored. Refer to the [Metadata Parsing Guide](#) for step-by-step instructions on parsing metadata.

### Uploading Metadata and Data to the NPS Data Store

All non-sensitive data sets should be uploaded to the NPS Data Store for wider dissemination. If the data set contains sensitive information the data set cannot be posted to the NPS Data Store but the data manager can choose to post the metadata file alone to the data store and maintain the data set locally. Refer to the [Uploading Data and Metadata](#) guidance for step-by-step instructions. Currently only the NCRN Data Manager and GIS Specialist have the appropriate permissions to upload Network data to the Data Store.

## Responsibilities

- Project Scientist (NPS staff or cooperator): The person creating or editing either geo or non-geospatial data is also responsible for FGDC compliant metadata development for any data set generated or modified as part of an NPS natural resource project. FGDC compliant metadata must accompany all geospatial datasets and is due when the final report and deliverables are submitted to the NCRN and parks included in the study. If you have not generated metadata in the past, or if you have any questions, please contact the data or project manager for assistance.

## Metadata Standard Operating Procedure

- Project Manager: The project manager must ensure that FGDC compliant metadata is received for all geospatial data received from the project scientist (Cooperator or NPS staff). The project manager should provide guidance and assistance to staff and cooperators seeking assistance with metadata creation.
- Data Manager: The data manager meets with the cooperators and/or NPS staff at the onset of a project and will develop a metadata 'template' document which has much of the standard information completed already. The responsibility for updating the metadata record belongs to the cooperator or project manager. The data manager is also responsible for reviewing the metadata and working with either the project manager or cooperator to make any necessary changes. Once a metadata record has been finalized, the data manager is responsible for posting that record to the NPS Data Store.

## Appendix A

1. Select the all of the parks where the project took place:

|                               |                               |                               |                               |
|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| <input type="checkbox"/> ANTI | <input type="checkbox"/> CATO | <input type="checkbox"/> CHOH | <input type="checkbox"/> GWMP |
| <input type="checkbox"/> HAFE | <input type="checkbox"/> MANA | <input type="checkbox"/> MONO | <input type="checkbox"/> NACE |
| <input type="checkbox"/> PRWI | <input type="checkbox"/> ROCR | <input type="checkbox"/> WOTR |                               |

2. What is the title of the data set?

3. Who are the creators/owners of the data set (include addresses, phone numbers and e-mail)?

a. If someone else should answer questions about the data, please list their name and contact information.

4. If the data set has been published please provide the citation information:

5. Please provide a brief description or abstract for the data set:

6. Please indicate the purpose behind the project and resulting data set:

7. Please indicate when the data was collected:

a. Single date:

b. Date range:

i. Start Date

ii. End Date

8. Status of the data set:

a. Is the project complete, in progress or planned?

b. Will the data set be updated?

i. If so, how frequently?

9. Please indicate appropriate keywords for the following categories:

a. Thematic

b. Place

c. Temporal

## Appendix A

d. Taxonomic

10. List any related data sets that could be documented for cross-reference:
  
11. Does the data set contain biological information? If no, skip to question 12.
  - a. What species or communities were studied?
  - b. Was a taxonomic authority or field guide used for identification?
  - c. Were voucher specimens collected?
    - i. If so, were they herbarium specimens, animal specimens, or photographs?
  
12. Was the data set developed using an analytical tool or model?
  - a. If yes, please reference the tool or model and provide a URL if one exists.
  
13. Describe the methodologies behind data collection and analysis:
  - a. Summarize field methods (copy/paste from other documents):
    - i. If existing protocols were used, provide the reference.
  
  - b. Summarize laboratory methods (if any):
    - i. If existing protocols were used, provide the reference.
  
14. Describe the quality assurance and quality control measures taken to ensure that the data are accurate:
  
  
15. Is this a GIS data set?
  - a. If so, how was the data set created (e.g. GPS, digitizing)?
  - b. Does the data set contain the proper projection information (e.g. projection, units, datum)
  - c. Provide a brief description of the data set:
  
  - d. List geo-processing steps:
  
16. Does the data set contain sensitive information? If so, describe:
  
17. For all data sets: does the data set contain sensitive information? If so, describe:
  
18. Provide any advice/caveats for potential data users: