

UC Santa Barbara

GIS Core Curriculum for Technical Programs (1997-1999)

Title

Unit 7: Finding, Creating, and Interpreting Metadata

Permalink

<https://escholarship.org/uc/item/3636p2fc>

Authors

7, CCTP
Larsgaard, Mary L.

Publication Date

1998

Peer reviewed

UNIT 7: FINDING, CREATING, AND INTERPRETING METADATA

Written by Mary Lynette Larsgaard, Map and Imagery Laboratory,
Davidson Library, University of California, Santa Barbara

Context

The quick definition of "metadata" is "data about data." Like most circular definitions, this is not very helpful. More usefully, metadata is relatively brief descriptive information about an item, that is a surrogate for that item, like a record in an online catalog in a library. For our purposes, metadata is surrogate information about digital spatial data. If you could directly search data, you might not need metadata; when this is not possible, then metadata is essential, if you want to search quickly through your own data and especially if ever you want other people to be able to find your data. It is essential also to enable users of GIS data easily to find out enough about data in a GIS so that the data may be easily processed and interpreted. In the context of GIS, metadata may exist at several different levels: for a GIS database as a whole; for individual layers; for individual tiles; and for individual points, lines, and polygons.

The purpose of metadata is to enable a user to find needed information quickly and efficiently. One difficulty with metadata is that cataloging-or, if you prefer, metaloging-is a considerable amount of work. This is because of the many standards that the library world uses to generate metadata:

- a set of cataloging rules (the Anglo-American Cataloguing Rules, AACR);
- a set of field order and punctuation rules (the International Standard for Bibliographic Description, e.g., ISBD(CM), for Cartographic Materials);
- several sets of subject heading lists of which the most prominent is the Library of Congress Subject Headings (LCSH); a database format for communication in online systems (USMARC = United States Machine-Readable Cataloging format); and the most current, the U.S. Federal Geographic Data Committees Content Standard for Digital Geospatial Metadata (FGDC; first edition, 1994; second edition in draft, 1997);

[outdated link removed]). On the other hand, given the type of fields that are available for cataloging GISs, the producer of the data is the best possible cataloger of the data.

The advent in 1994 of FGDC has been extremely influential in the world-wide application of metadata to digital geospatial data. An excellent Website providing links to the ongoing metadata efforts in this field is "Jan Smits' bookmarks concerning metadata for digital spatial data," *[outdated link removed]*. This provides links to international efforts such as ISO/TC211 and CEN/TC287, and to national-level efforts, such as the British DESIRE. For persons in the U.S., two useful pages are *[outdated links removed]*

The former is the page of the U.S. government agency that leads metadata efforts within the Federal government; the latter is the homepage of the Alexandria Digital Library, whose goal is to provide online access to georeferenced information.

A portion of the FGDC page of especial interest is the "Metadata Tools Survey, July 1997" *[outdated link removed]* which lists and briefly annotates about fifteen different tools for generating metadata. A prominent example of automatic extraction of metadata is ESRI's **document.aml**. Document.aml (version 7.1.1. to be the last while ESRI works on new metadata documentation tools) is available from ESRI's Web site at <http://www.esri.com>; go to the site and do a search on the term "document.aml" for instructions on how to get a copy of the software and steps in using it. For an article with examples of using document.aml, read "Automated Capture of Metadata; Simple Procedures and Tools for Editing Coverages" by Dan Williams, go to *[outdated link removed]*

An effort to make manual metadata logging easy by using a relatively limited number of fields is the Dublin Core. The Dublin Core contains fifteen fields:

1. subject
2. title
3. creator/originator
4. publisher
5. other agent/contributor
6. date
7. resource type
8. format
9. resource identifier
10. source
11. language
12. coverage
13. rights management
14. relation
15. description

These metadata fields are intended to be the minimum amount of information that a Webpage producer can enter in order to assist prospective users to find the page. The Dublin Core homepage is at http://www.purl.org/metadata/dublin_core. Several agencies have used either

Dublin Core or some more limited subset of it, and in some cases have homepages where you may catalog your homepage by filling in a form. The Nordic Metadata Project's "Dublin Core Metadata Template" is an example of this. It is at *[outdated link removed]*. For examples of others, see the Dublin Core generator of the United Kingdom Office for Library and Information Networking (UKOLN), at *[outdated link removed]*, or one from Australia at *[outdated link removed]*

If you feel brave or if you want maximum documentation of your GIS data, use the FGDC document *Instructions for Implementing the FGDC Metadata Standard, and Testing Conformance* (text and word-processor versions are at: *[outdated link removed]*). The FGDC document has the following sections:

1. Identification information;
2. Data quality information;
3. Spatial data organization information;
4. Spatial reference information;
5. Entity and attribute information;
6. Distribution information;
7. Metadata reference information;
8. Citation information;
9. Time period information;
10. Contact information.

Citation information forms the basic bibliographic reference. Fields in the latter and in *Time period information* and *Contact information* may be used as often as they are needed. Section 5, *Entity and attribute information*, is specifically for documenting GIS data.

Encouragingly enough, there are occasionally workshops to teach metadata creation. See for example the national satellite videoconference, "A Practical Guide to Metadata Implementation for GIS/LIS Professionals," Wednesday, October 15, 1997, 1-3pm CDT. This conference--produced by the University of Wisconsin, Cooperative Extension--has the following sections: What is Metadata and Why Is It Important?; Get Acquainted with the Content Standards for Digital Geospatial Data [FGDC]; Getting Started; Metadata Tools; Practical Experience with Creating Metadata, a State Agency Perspective; Other Metadata Issues.

There is another Federal effort to document data, the **Government Information Locator Service** (GILS). The object of GILS is to enable potential users of Federal information to readily locate that information. For a report on GILS, see:

Moen, William E.; and McClure, Charles R. 1997. An evaluation of the Federal government's implementation of the Government Information Locator Service (GILS). *[outdated link removed]*

Example Application

1. You need to find existing information in digital form for a GIS database that you are compiling, and you don't want to scan or digitize hardcopy data that is already available in digital form. Use Webbrowser searches, or, for a short cut,
 2. You are looking at metadata for a digital data set; what does all this gobbledygook mean? If it is in FGDC form, to find out the meaning of the fields, go to <http://www.fgdc.gov>. If it is from a library's online catalog, consult a cataloger, or read AACR (current version) or USMARC.
 3. You need to use an existing set of digital data. But before you use the data, you need to know about data quality - accuracy, lineage, source of information. Go to the FGDC standard (in Website given in 2. above), and read section 2, "Data Quality Information." Take a look at the metadata for AVHRR, [outdated link removed], or any other metadata you can find. Figure out what all that information really means, and what it tells you about the quality of the data - is this a good dataset or not, and why or why not?
-

Learning Outcomes

The following list describes the expected skills which students should master for each level of training, i.e. Awareness/Competency/Mastery.

Awareness:

The expected learning goals of this section are to achieve a general understanding of metadata, its potential uses, and a working knowledge of vocabulary.

Competency:

The learning goals of this section are to develop the ability to find, use, and interpret complicated metadata, and to create simple metadata.

Mastery:

The learning goals of this section are to be able to understand the intricacies of complicated metadata, and to efficiently create, navigate, and catalog it.

Preparatory Units

Recommended

- [some from below?]

Complementary

1. (UNIT 1) - Data acquisition
 2. (UNIT 31) - Managing database files
 3. (UNIT 32) - Managing digital libraries
 4. (UNIT 52) - Project management
-

Awareness

1. Student can use online catalog to find metadata. Either use a local library's online catalog, or go to: *[outdated links removed]*. Example task: search for "AVHRR".
2. Student can use Web search engines to find metadata. Use different search engines to look for "AVHRR" and "AVHRR metadata". See which engine(s) works best.
3. Student knows basic digital data description. Read through Dublin Core homepage.
4. Student knows basics of GIS description. Read FGDC metadata document, concentrating on Section 5.

Vocabulary:

- AACR
- AVHRR
- DCW
- DLG
- Document.aml
- Dublin Core
- FGDC
- Georeferenced information
- Geospatial information
- GILS
- LCSH
- Metadata
- Online catalog
- spatial data

Topics/Generic List of Tasks

1. Go to MELVYL or to your local online catalog. Search for any of the following: AVHRR; Digital chart of the world (DCW); Digital line graph (DLG); etc.
2. Using several different Web search engines, try to find metadata for the same items. Also, search the National Geospatial Data Clearinghouse (enter this at <http://www.fgdc.gov/>).
3. Read through the Dublin Core; determine which fields are most useful for describing a GIS database.
4. Read through the FGDC metadata document (*[outdated link removed]*). Determine what specific fields besides those in Section 5 are of most use for describing a GIS database.

Competency

Learning objectives:

1. Student will find extensive metadata for GIS data.
2. Student will compare and contrast metadata for different types of digital geospatial data.
3. Student will do Dublin-Core metaloging of own Webpage or of a GIS dataset as a whole. Use the Nordic Metadata form, or other form available over the Web.

Generic list of tasks:

1. Select subjects of interest to you, and then try to find metadata for that type of data. Or use the following metadata for AVHRR: *[outdated link removed]*
 2. Compare this with the record you found for any AVHRR dataset that you found on MELVYL.
 3. Go to the Nordic Metadata workform and catalog your Webpage or GIS dataset. Download the metadata to your computer.
-

Mastery

Learning objectives:

1. Search extensively on the Web and try to find all the digital geospatial data you can locate on an area and subject of your choice.
2. Create full metadata for your, or a sample, GIS database.
3. Explain what it all means to someone who knows nothing about it.

Generic list of tasks:

1. Select the search engine that you found to be most useful in earlier exercises. Pick a subject or region and compile an extensive list of metadata records for your area of interest.
 2. Call up the FGDC standard. Download it. Fill in the fields for your GIS dataset.
 3. Find a person who understands GIS but knows nothing about metadata. Explain every field in the metadata.
-

Follow-up Units

- (UNIT 53) - Communicating about and distributing GIS products
-

Resources

- Anglo-American cataloguing rules, second edition revised. 1988. Chicago: American Library Association.
- USMARC concise formats. 1991- . Washington, D.C.: Library of Congress, network Development and MARC Standards Office.

Created: May 14, 1997. Last updated: October 5, 1998.