Metadata Primer for Map Librarians

Electronic Publication No. 3

(This article was largely written by David Y. Allen. The author’s role has consisted of reworking the ideas of others, who are mostly responsible for its intellectual content. Mary Larsgaard, Patrick McGlamery, and Jan Smits contributed substantially to this article and made numerous suggestions and corrections. They are not responsible for any errors.)

Introduction

Confused about Metadata? Not to worry--so is everybody else. Metadata is undoubtedly one of the most complex issues facing map librarians in the digital age. There are several reasons for the confusion. First, there are a number of competing standards for metadata. These standards are evolving and, in some cases, converging. To further complicate matters, there are many different types of digital files that may be described by metadata. A type of metadata that works well with one kind of file may be too cumbersome or inadequate to describe others.

The situation becomes more manageable if one is concerned about producing metadata for a particular application. None of the metadata standards are that difficult to understand and apply. Our purpose here is to provide an overview of existing standards, to show you where to get more information, and to suggest which types of metadata may be most appropriate for particular kinds of digital files.

What is Metadata?

The term "metadata" originated among GIS specialists when they realized they needed detailed information about the computer readable files they were producing. The classic definition of metadata is "data about data." A more helpful one is "information about a publication, as compared with the content of the publication". Metadata is generally taken to mean information about digital data only, although as is evident from the definitions, that is more often implied than explicit.

Metadata resembles cataloging, but it was not designed to be produced by librarians, but rather by the producers of the data--i.e. the creators of the digital files. Most metadata is designed to be indexed and retrieved on the Web, rather than to be retrieved on conventional online catalogs.

What is the difference between metadata and cataloging information? At this point, very little. Some librarians have suggested that the advocates of metadata are reinventing the wheel, and maintain that standard MARC cataloging can be adapted to do everything that metadata does. Nonetheless, there is no consensus even among map catalogers on this issue. Good metadata should include all of the information needed to read and interpret a complex digital file. For example, metadata should also be able to include graphic fields (that is, not just URLs linking to graphics, but the actual graphics themselves). MARC is not designed to convey complex numeric information, and can be adapted much more easily to produce records for simple map images than complex GIS data sets.

Rather than attempt to decide on a single metadata standard, we would suggest that various types of metadata may be appropriate for various file types, and that librarians in different situations may need to work with one or more kinds of metadata. What is perhaps most important is to maintain as much
consistency between various types of metadata as possible. If that is done, records can readily be interchanged using "crosswalks," and can be displayed in a single database.

Types of Metadata

Numerous metadata standards have been developed. Leaving aside for the moment international standards, three types of metadata are dominant in the United States.

**FGDC Metadata:** In the United States the most established metadata standard is that adopted by the Federal Geographic Data Committee (FGDC). This standard is used by the federal government and many other agencies. In fact, it is obligatory for federal agencies to produce FGDC metadata for their digital products. Detailed information about FGDC metadata is available at the FGDC home page (http://www.fgdc.gov/metadata). Here can be found tutorials, examples, and all the other information you need to get started making FGDC metadata. Probably more metadata can be found in FGDC format than in any other. A major effort is underway to harmonize the FGDC standard with European standards, and a new international standard with the blessings of the International Standards Organization (ISO) will soon replace the FGDC standard.

Although the FGDC is clearly winning widespread acceptance, it is not without problems. The most common criticism of FGDC metadata is that it is too cumbersome and difficult to apply—that it is overkill for very simple digital files. To address this problem, a number of states have adopted simplified versions of FGDC metadata ("metadata lite"). The FGDC itself is developing a list of essential metadata elements—its own version of "metadata lite."

Not the least of the problems with FGDC metadata is its incompatibility with MARC. You can't find FGDC metadata on OCLC or RLIN. On the other hand, the FGDC records are more extensively indexed than MARC, and can be readily searched on the World Wide Web.

Finally, even FGDC metadata is not sophisticated enough to suit everyone's needs. Although it is more information rich than MARC, it is inadequate for those attempting to construct graphical interfaces (like Project Alexandria or Harvard's "Liboratory"), and needs to be supplemented for these purposes. Clickable maps, which allow a person to retrieve digital files for a particular area, are an increasingly popular option not provided for by the FGDC standard.

**MARC as Metadata:** There are obvious advantages to having a single standard for cataloging both digital and non-digital materials—hence the appeal of MARC. Recognizing that traditional MARC records are inadequate for many digital materials, librarians have been attempting to upgrade MARC to incorporate additional data elements from the FGDC. For further information on cataloging cartographic datafiles in USMARC, see "Guidelines for Distinguishing Cartographic Records on Computer File Carriers from other Materials on Computer File Carriers," (http://www.lcweb.loc.gov/marc/cfmap.html) which has been prepared by the Library of Congress.

**Dublin Core Metadata:** The "Dublin Core" bears a relationship to USMARC, which somewhat resembles that between "metadata lite" and FGDC metadata. The Dublin Core is named after Dublin, Ohio, the home of OCLC, which is the chief proponent of this metadata standard. The authoritative source for information on this standard is the OCLC Dublin Core Homepage (http://purl.oclc.org/dc/). The Dublin Core has been rapidly evolving, and remains the subject of much discussion and controversy. In this regard see the Final Report of the MAGERT CCC Task Force on Using Dublin Core for Cartographic Materials (/magert/committeesabc/catalogingandclassification/dcfinalrpt).
In the opinion of some of us, the Dublin Core has potential as a form of brief cataloging for maps, including raster images of maps available on the Internet. The Dublin Core consists of fifteen "core" fields, which more or less correspond to certain fields in MARC records. The Dublin Core is extremely flexible: all of the fields are optional, and all are repeatable. Once a basic template is devised for maps, it should be possible for a non-professional with little training to produce acceptable Dublin Core records for cartographic materials in both digital and non-digital form. Dublin Core records can be searched on the Web, and experiments are being made by OCLC to produce databases that utilize both Dublin Core and USMARC records. The important OCLC CORC Project (http://purl.oclc.org/corc/) contains many examples of map cataloging records that can be displayed in both MARC and Dublin Core formats.

Levels of Metadata

Implicit in the above discussion is the idea that "one size fits all" does not work with metadata--that no single standard is suitable for all purposes. The various types of metadata can be placed in a hierarchy of complexity. Jan Smits has discussed in some detail the need for various levels of metadata in "The Creation and Integration of Metadata in Spatial Map Collections" (http://magic.lib.uconn.edu/ifla/meta-smits.htm) (Link no longer active), and those interested in pursuing the subject in depth should consult his article. In summary, those who are trying to describe complex GIS data sets will probably need to work with the FGDC/ISO metadata. MARC records can be used with less complex data sets. And Dublin Core, as well as MARC, is suitable for raster images and simple vector data sets that do not require a lot of description.

An important point is the need for consistency and linking among the various levels of metadata. Very often map librarians may want to link FGDC metadata with MARC or Dublin Core records providing briefer descriptions of the same data (for distribution on online catalogs, etc.). If this is done, the URL for the FGDC metadata should be provided in the MARC record. It is also important to provide as much consistency as possible between records in different formats, as is most obviously the case with name authority.

Another reason for keeping the fields in different metadata records as consistent as possible is to facilitate the interchange of records. Thus, at some point you may want to upgrade your Dublin Core records to full MARC cataloging. Some librarians receiving grants from the federal government have found it necessary to translate their metadata in MARC format into the FGDC standard. And those with FGDC metadata on their hands may need to create MARC records for distribution via online catalogs.

To facilitate moving metadata from one level to another, a number of so-called "crosswalks" have been developed. A good collection of these is Mapping between Metadata Formats (http://ukoln.ac.uk/metadata/interoperability/) from UKOLN, the UK Office for Library and Information Networking. A FGDC to Marc Crosswalk (http://www.alexandria.ucsb.edu/public-documents/metadata/fgdc2marc.html) is available from the Project Alexandria site. Most recently, Adam Chandler and Dan Foley at the U.S.G.S. Energy & Environmental Resources Center have written a paper describing their project to develop a (link no longer active) crosswalk between Dublin Core, USMARC and FGDC (http://eeirc.nwrc.gov/pubs/crosswalk/fgdc-marc-dc.htm).

Other Sources of Information

Aspiring Masters of Metadata can explore the subject in depth at the following Web sites:

IFLA Metadata Resources Page (http://www.ifla.org/II/metadata.htm)

Metadata Page (http://ukoln.ac.uk/metadata/) (from UKOLN, the UK Office for Library and Information Networking)
Project Alexandria metadata page (http://www.alexandria.ucsb.edu/frames3.html) (link no longer active)

**An important source of information about metadata for maps in paper:**

Paige Andrew and Mary Lynette Larsgaard, eds. *Maps and Related Cartographic Materials: Cataloging, Classification, and Bibliographic Control* (New York: Haworth Press, 1999). This book, which was co-published simultaneously as v. 27, no. 1/2 and 3/4 of *Cataloging and Classification Quarterly*, contains articles discussing in depth most of the issues mentioned here.

**Keep up with current developments with these online publications:**

- d-lib Magazine (http://www.dlib.org/)
- RLG DigiNews (http://www.rlg.org/preserv/diginews/index.html)

*Revised February 12, 2001*