
Using the Metadata Object Description Schema (MODS) for resource description: guidelines and applications

Rebecca S. Guenther

The authors

Rebecca Guenther is Senior Networking and Standards Specialist, Network Development and MARC Standards Office of the Library of Congress, Washington, Dc, USA.

Keywords

Online cataloguing, Standards, Archives

Abstract

This paper describes the Metadata Object Description Schema (MODS), its accompanying documentation and some of its applications. It reviews the MODS user guidelines provided by the Library of Congress and how they enable a user of the schema to consistently apply MODS as a metadata scheme. Because the schema itself could not fully document appropriate usage, the guidelines provide element definitions, history, relationships to other elements, usage conventions, and examples. Short descriptions of some MODS applications are given and a more detailed discussion of its use in the Library of Congress' Minerva project for Web archiving is given.

Electronic access

The Emerald Research Register for this journal is available at www.emeraldinsight.com/researchregister

The current issue and full text archive of this journal is available at www.emeraldinsight.com/0737-8831.htm

The need for a rich XML-based descriptive metadata standard such as the Metadata Object Description Schema (MODS) has been expressed by members of the digital library and related communities as they attempt to implement projects involving search and retrieval, management of complex digital objects, integrating metadata from library databases with other non-MARC sources, and other functions. There are several applications that have a metadata component for which it is desirable to have richer metadata than offered by the Dublin Core Metadata Element Set, but that is less complex and more user friendly than full MARC 21 with its numeric tags (Dublin Core Initiative, 2003). Sally McCallum discusses some of these in her article in this issue.

Users of MODS benefit from the energetic and systematic support of the Library of Congress staff and others who have developed the MODS fields based on vast and broad experiences. LC has supplied detailed documentation, examples of usage, and other supporting materials such as guidelines and local policy to facilitate the use and adoption of MODS. It will remain well supported, maintained and documented because of the commitment of the Library of Congress to standards development and the contributions of outside participants.

MODS user guidelines

Need

In order to use a metadata element set, particularly for the creation of new records, it is necessary to have both the available elements and their meanings (semantics), a method for encoding them (in the case of XML, a schema), a system that allows for input, and ultimately a system that can effectively use them. The Network Development and MARC Standards Office of the Library of Congress developed MODS initially as a MARC derivative, both as an element set (essentially a subset of MARC with different element names) and as a syntax

Received 12 September 2003

Revised 13 October 2003

Accepted 7 November 2003

© Rebecca Guenther



(in XML using element names). It became clear early on that the schema itself could not fully describe each element and guide users to its application. Hints in the schema are given using the documentation tag, which points the user to the appropriate equivalent MARC element. This, however, was not sufficient for applying MODS in many cases where there was not a clear mapping or more information on usage was required. LC has therefore provided extensive guidelines.

General overview

The MODS user guidelines attempt to provide the user with element definitions, history, relationships to other elements, usage conventions, and examples (Library of Congress, n.d., a). The guidelines are primarily intended to be used for assistance in creating original MODS records, although they may also be instructive for interpreting MODS records that have been converted from MARC 21 or for use in developing detailed conversion specifications.

Definitions are often in the form of links to MARC documentation, where MODS elements inherit MARC semantics. A mapping may be provided as well as a reference to a definition. The guidelines point out when an element is used in a different way than in MARC 21.

Introduction and implementation

The guidelines include a basic review of XML structures with definitions and examples. The implementation section includes the relation of MODS to cataloging rules and full MARC 21, how to handle punctuation, information about general element repeatability and order, and instructions on encoding multiple MODS records versus a single MODS record.

Punctuation

The MODS guidelines specify that punctuation is basically generated based on the particular element (for instance ISBD punctuation), except where it occurs within an element. This is accomplished with an XSLT stylesheet, which inserts the required punctuation on the basis of the tag. However, since there will be cases when a MARC 21 record is converted to MODS, it may not always be possible to extract

all ISBD punctuation, and it may be retained in the data after conversion. For new MODS records, the principle should be followed that punctuation between elements is generated by a stylesheet if ISBD display is desired.

Attributes used throughout the schema

There are a number of attributes that are referenced from different elements throughout the MODS schema. Because XML schema allows for complex types that can be reused, this brings a high level of consistency to the element definitions. (There are numerous places in MARC that define similar elements the same way, although constraints inherited in the limited number of subfields that may be included does not always allow for full consistency.) Many of the attributes used widely throughout the schema concern language issues regarding the metadata. The schema allows for attributes for each top level element in MODS to indicate the language, script and transliteration scheme of the metadata in the given element. The guidelines elaborate on how these attributes should be used, and points out when they do not exist at certain levels in MARC 21.

Detailed description of elements

The guidelines describe each element (including its subelements) and attributes in detail. It provides links to the MARC 21 documentation for equivalent elements, specifies which MARC 21 fields are equivalents and which fields are collapsed to one MODS element (since MODS is a subset of MARC), indicates under what conditions an element is repeated, and gives recommendations on best practice. The guidelines are particularly useful for elements that do not exist in MARC 21 or do not have a direct equivalent, since in these cases there is generally no other documentation to reference definitions. For instance, the element `<displayForm>` does not have a direct equivalent in MARC. It is related to `<245$c>` (statement of responsibility), although this mapping is misleading, since that MARC subfield often carries other data than merely the display form of a name. (For this reason, the MARC to MODS mapping uses a note with a type attribute "statement of responsibility".) The entry for `<displayForm>` is as follows:

“displayForm” is used to indicate the unstructured form of the name. It includes a name as given on the resource. For some applications, contractual agreements require that the form of the name given on the resource be provided. This data is usually carried in MARC 21 field 245 subfield \$c, although the latter may include other information in addition to the display form of the name. If part of an entire statement of responsibility, it may be indicated in a note (with type=“statement of responsibility”) along with any other text in the statement of responsibility and also repeated here if desired.

Many elements are packaged differently in MODS than they are in MARC 21. For instance, the MARC Leader/07 (bibliographic level) includes several different concepts, including issuance (whether serial or monographic), granularity (whether a collection or single item), and method of update (whether continually updated or static). In MODS an element <issuance> is used under <originInfo> with enumerated values “monographic” or “continuing.” Whether or not the item is part of a collection is indicated as an attribute associated with <typeOfResource>, since any of the enumerated high level resource types may be collections or single items (collection=“yes” indicates that the MODS record describes a collection). In addition, MODS uses the attribute manuscript with a value of “yes” or “no” in <typeOfResource> to indicate whether the high level resource type is in manuscript form (MARC has separate Leader/06 (type of resource) values for these). The guidelines relate these elements and attributes to the appropriate MARC element and describe how they are to be used.

Because MODS was developed particularly with electronic resources in mind, there are a number of elements that give information specific to these types of resources. For instance, under <physicalDescription> there is an element <digitalOrigin> which does not exist in MARC 21, but was felt to be useful for description of electronic resources. It designates whether the item is born digital or reformatted digital. The definition is as follows:

“digitalOrigin” designates whether a digital resource is born digital or reformatted digital. It does not exist in the MARC 21 formats, however, it is beneficial in the description of digital materials. The following values may be used:

born digital – A resource was created and is intended to remain in digital form.

reformatted digital – A resource was created in a non-digital form and was converted into a digital form.

A few specific MODS elements are discussed below in order to illustrate the use of the guidelines.

MODS notes

In order to accommodate a wide range of types of note fields by various different applications, the MODS team at LC decided to allow for an open-ended list of MODS types. To enhance the likelihood of interoperability between applications, the Library of Congress (n.d., b) has made available a list of note types used by implementers, which is referenced in the guidelines. Much of the input to the list at the time of the writing of this article is from the University of California at Berkeley (n.d.), which used its own schema in their database for digital objects, generic descriptive metadata (GDM), but is converting to MODS. The list will be continuously updated as users submit note types that have been input into their MODS records. Users are free to use any value for the note type attribute regardless of whether their types have been submitted to LC.

Related item

The <relatedItem> element is a powerful feature in MODS in that it allows for rich linking and description of related resources. The MODS guidelines are essential in describing the intent and features of this element because of its complexity. <relatedItem> includes an attribute “type” to indicate the type of relationship between the resource described in the body of the MODS record and the resource described in <relatedItem>.

The guidelines explain that there are several MARC 21 fields that are mapped to <relatedItem>. These include the 7XX added entry fields, but only those that contain a title subfield (subfield \$t), since the existence of a title subfield denotes an analytical author/title entry (i.e. the work or part of the work is contained in the resource). Thus, in these situations the type attribute has a value of “constituent” (i.e. contained in the work

described). In addition, all the linking entry fields (76X-78X) map to <relatedItem>; in some cases multiple MARC linking entry fields are mapped to the same <relatedItem> type. For example, field 770 (supplement/special issue entry) and field 774 (constituent unit entry) both are mapped to type="constituent", since 770 denotes a supplement contained in the whole, which is a special type of constituent. It was felt that this distinction was not necessary to retain in MODS. Series was considered yet another relationship between resources, so type="series" is used for series statements, mapped to various MARC fields that include series information (440, 490, 760, 762, 800, 810, 811, 830).

The <relatedItem> element is fully recursive so that any MODS element may be used. This allows for rich description and the ability to designate hierarchical relationships. The guidelines elaborate that full description for constituent items (using type="constituent" and any MODS element) may be included, which might be suitable for complex digital objects that require specific descriptive information for constituent parts, but that are considered intellectually one object (e.g. a CD with several tracks, digitized or analog). Some of the elements available in MODS in <relatedItem> are not available in the MARC linking entry fields, and further development of the latter is limited by available subfields, which is restricted to using numerics 0-9 and one-character alphabetic subfield tags.

In addition to all the defined MODS elements, in MODS version 3.0 <relatedItem> includes elements for parsing the details of citation numbering, to be used when the MODS record describes an article. This will enable the construction of an OpenURL or a formatted citation.

Identifier

The identifier contains a standard number or code that identifies the resource. The guidelines explain that there is no controlled list of values for the type attribute, which designates the identifier scheme used for formulating the identifier recorded. It gives a list of suggested values for identifier scheme, however, and states that others may also be used. If the user desires to record a local identifier either the type

value "local" may be used, or the specific named identifier scheme may be recorded, since this is not a controlled list.

Extension

Because the extension element has no substructure in the MODS schema and no clear equivalent in MARC (although it can be mapped to 9XX), there is little guidance in the schema on how to use it. The MODS guidelines describes situations that warrant its use and gives some examples. The element may be used for local elements, similar to MARC 21 9XX fields, that the user wants to include but are undefined in MODS. In addition, it is a mechanism for including elements from other XML schemas that may enhance the MODS record.

MODS Lite

Although MODS includes fewer data elements than MARC 21, one can look at the schema as fairly complex. Allowing for a simpler expression of MODS results in better crosswalking to other simple schemas, such as Dublin Core. Various transformations may be possible using stylesheets and software tools developed at the Library of Congress' Network Development and MARC Standards Office, which is part of the framework for working with MARC data in an XML environment. Corey Keith's article in this issue details MARCXML developments.

A "MODS Lite" is described in the MODS guidelines to show how a simple MODS expression might be used. It shows how to express metadata using only MODS elements that are equivalent to the 15 elements of the simple Dublin Core Metadata Element Set. The substructure defined in the MODS schema is used to allow for records that validate, but elements that do not have a direct equivalent in Dublin Core are not included (except that record information is given in MODS Lite but not in Dublin Core). Various subsets of elements from MODS could define other MODS "Lites"; the Minerva project detailed below uses a different subset. In the future, the Library of Congress will publish other MODS Lites that can be used for other applications.

Additional features

The guidelines include full record examples to assist the user in seeing the relationship between MODS elements. It is expected that over time this section will grow considerably as more implementations contribute full MODS record examples. In addition, there is an index by element name regardless of its hierarchy in the element list to facilitate finding information quickly within the guidelines. Of course, users of MODS Lite can extend the structure to include other elements that they feel are essential, maintaining consistency with both MODS and Dublin Core.

Applications

MODS has been gaining some attention as a descriptive metadata scheme of particular interest for digital materials. The existence of both tools for transforming records between schemes and extensive guidelines should facilitate its further adoption. This section will describe some applications of the MODS schema.

Although this article will focus on the application of MODS to the Minerva project (Mapping the INternet Electronic Resources Virtual Archive) it is worth noting some additional MODS metadata projects (Library of Congress, n.d., c). A number of these are METS projects that use MODS as the descriptive metadata extension schema. This list is not comprehensive, but illustrates some applications of MODS.

- The University of Chicago Library (Olson, 2003) is crosswalking MARC to MODS for use as descriptive metadata embedded in METS objects for one of its digital library collections. In this case, all of the electronic objects were created from physical objects in their collection which already had detailed cataloging records. (MODS list 26 June 2003.)
- LC Digital Audio-visual Preservation Prototyping Project (Library of Congress, n.d., d) explores aspects of digital preservation for audio and video. It packages digital content and metadata using METS. Descriptions of objects and subobjects use MODS. Where previously cataloged descriptions exist in the Library of Congress catalog, the MARC records are transformed to MODS and loaded into the database used. Where they are not available, original MODS records are created.
- OAI Harvesting. The Library of Congress incorporates MODS as an alternate format for its over 100,000 metadata records for items digitized for American Memory that are harvested (items represented include books, maps, photographs, early movies, sheet music, and printed ephemera) (Library of Congress, n.d., e).
- University of California Bancroft Library is using an SQL Server to gather descriptive metadata to be used in the context of METS objects. The database, originally designed around the MOA2.DTD, gathers all of the descriptive, administrative and structural metadata necessary to generate METS objects with MODS descriptive metadata and (minimal) MIX image technical metadata. Although the descriptive metadata were not designed around MODS but long before it existed, the elements have been mapped to MODS for use as descriptive metadata (Beaubien, 2003).
- Some users are experimenting with MODS for bibliographies and generally bibliographic citations. Version 3.0, which at this writing is undergoing review, adds additional elements to `<relatedItem>` to accommodate parsed citation information that will assist in coding bibliographic citations and generating an OpenURL link.
- The National Library of Australia is using MODS for several projects. It is being used as an exchange format to facilitate harvesting and conversion of data in multiple original formats/schemas to MARC and subsequent loading into Australia's National Bibliographic database (NBD). This allows NLA to include records from non-MARC, non-traditional contributors in its union catalog by loading them after converting by using XSLT stylesheets. In addition MODS will be used as a storage format for the MusicAustralia

service, which involves gathering data from various contributors – libraries, archives, museums, the commercial sector – via the Harvester project, and loading all records into the NBD. NLA extracts the records it wants from the NBD and converts them to MODS for the MusicAustralia service itself (Ayres, 2003).

MINERVA project

MINERVA is an experimental pilot developed to identify, select, collect and preserve open-access materials from the World Wide Web. The effort includes consensus building within the library; joint planning with external bodies; studies of the technical, copyright and policy issues; the development of a long-term plan; and coordination of prototypes. A multidisciplinary team of library staff representing cataloging, standards, legal, public services, and technology services is studying methods to evaluate, select, collect, catalog, provide access to, and preserve these materials for future generations of researchers

The project is a collaboration between the Library of Congress, the Internet Archive (Alexa), State University of New York (SUNY), and the University of Washington. The latter groups have assisted in identifying content and in using tools of their design to assign metadata descriptions to the Web sites collected. This metadata database is then used to search, retrieve and analyze the archived collection of Web sites. The contractors collected and archived Web sites based on LC specifications focused on themes concerning the presidential Campaign 2000; 11 September, 2001; Winter Olympics 2002; Election 2002; and the 107th Congress.

Metadata is created for Web sites in the collection using the MODS schema because of its compatibility with MARC data. The data in the MODS records is used in the search and retrieval system for Minerva. The library is also beginning to experiment with METS to enable the inclusion of additional metadata, such as technical, provenance, and rights metadata. An important part of this work involves defining a METS application profile for Web sites. To support the project, LC's Network

Development and MARC Standards Office is currently working on tools for the creation of MODS records and the conversion of MODS to MARC so that the records can be brought into LC's online catalog.

The MINERVA collection on the theme of Election 2002 is the first to provide full metadata records for each of the sites collected. It is a selective collection of nearly 4,000 sites archived between 1 July, 2002 and 30 November, 2002. The Election 2002 Web archive includes Web sites produced by congressional and gubernatorial candidates as well as party, interest group, press, government, civic, and other selective Web sites related to the 2002 national and statewide elections. For MINERVA collections, the Library of Congress creates collection-level cataloging records that are accessible in LC's integrated library system. This is being supplemented by records for each Web site within the collection (one record per "base URL", which includes multiple snapshots of different dates captured) which are used within the Minerva Web archive search system. LC decided to use MODS for its record format because of its flexibility in level of markup, its compatibility with existing MARC records in the ILS, and its usefulness as a descriptive metadata extension schema with a METS object in a future repository.

Metadata preparation for Election 2002 Web Archive

The contractors used specifications supplied by the Library of Congress concerning which MODS elements to use and how to populate them. A small number of MODS elements are used for the resource descriptions. Following are the fields included along with descriptions. An XSLT stylesheet is used to present the data in the form of a catalog record that is easy to understand.

- *Title.* Title data is taken from text included within the title tag of the HTML source file of the resource at the URL being described. In addition, contractors constructed an alternative title based on information about the site such as candidate name, party, office, state (or city), since HTML titles may not always reflect the true content. If the HTML title was not unavailable, they used this constructed title as main title. LC

- gave the contractors specifications for what data to include for each type of Web site. Categories include: candidate, citizen, civic and advocacy, government, political party, press, public opinion and miscellaneous.
- *Name.* A name is defined as a personal or corporate entity related to the resource; it may be the creator or issuing publisher. Contractors were instructed to include the name of the entity that appears to be primarily responsible for making the content of a Web page, as identified by reference to text or graphics on the captured page, or by reference to an “about us” page linked from the page. Name is encoded in a structured form, i.e. lastname, firstname and type of name is indicated (personal, corporate). Names are not authoritative in that they have not been validated against an authority file.
 - *Abstract.* A brief description of the site associated with a Web page, is given, generally generated from data collected about the site, referencing a possible site producer and identifying a possible purpose of the site. If available from a meta tag “description” within the source, it may be extracted.
 - *Date captured.* This is the archived time associated with the Web page which was archived between 1 July, 2002 and 30 November, 2002. This may be a range, with start and end dates (MODS allows for an attribute associated with the date to indicate whether it is a start or end date). The date of the first iteration and the date of the last iteration for each URL is included (Web sites were captured periodically over this period of time, so each has many iterations).
 - *Genre.* Each URL is identified as “Web site”. This data is generated since it is always the same.
 - *Physical description/format.* Contractors were instructed to provide a list of the distinct file formats expressed as standard Internet media types (e.g. text/HTML, image/jpeg) of all archived objects associated with a Web page.
 - *Related item.* The collection name of which the resource is a part (with type=“host”), e.g. Election 2002 Web archive and its URL is included. This provides a link to the collection Web site as well as its metadata. This field is generated, since the data is the same for each record in the collection.
 - *Identifier.* Two identifiers are given. One is the URL of the site which was collected, used with a display label “Active site (if available)”. Because it is possible that a site could have disappeared since the election (and indeed many have), a hot link is not provided for this URL. A second identifier is also included for the location of the archived site with an attribute displayLabel “archived site.” A hot link is provided for the archived site URL associated with the dates of capture, which allows the user to go directly to the archive that contains the captured files.
 - *Language.* The primary language of a Web page is identified using the ISO 639-2 bibliographic code (e.g. eng, fre). The XSLT stylesheet that generates the displayed record converts the code into text for the user display.
 - *Access condition.* An identifier or statement of access terms and conditions provided by the library is given if there are access or use restrictions associated with a Web page. Otherwise the text “none” is generated.
 - *Subject.* Library of Congress Subject Headings (LCSH) are used based on guidelines provided. In many cases these are generated using rules for each Web site type provided by LC and based on data already collected about each site. Some sites, such as citizen groups, do not have controlled subject headings, because it was not possible to provide consistent rules for supplying these. There are also cases where the rules were complex and the instructions resulted in headings that were not completely consistent with LCSH; these will be converted later, since they are predictable situations (an example is the use of postal codes for states instead of abbreviations specified in LCSH). If available from metadata in HTML tags keywords may be extracted and included as uncontrolled subject terms.

Figures 1-3 show an example of the XML coding for one Web site from the Election 2002

Figure 1 MODS document for candidate Fran Ulmer Web site

```

<?xml version="1.0" encoding="UTF-8"?><mods xmlns="http://www.loc.gov/mods/"
xmlns:xlink="http://www.w3.org/TR/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.loc.gov/mods/
http://www.loc.gov/standards/mods/mods.xsd">
  <titleInfo>
    <title>FranUlmer.com -- Home Page</title>
  </titleInfo>
  <titleInfo type="alternative">
    <title>Fran Ulmer, Democratic candidate for Governor, Alaska, 2002</title>
  </titleInfo>
  <name type="personal">
    <namePart>Ulmer, Fran</namePart>
  </name>
  <genre>Web site</genre>
  <originInfo>
    <dateCaptured point="start" encoding="iso8601">20020702</dateCaptured>
    <dateCaptured point="end" encoding="iso8601">20021203</dateCaptured>
  </originInfo>
  <language authority="iso639-2b">eng</language>
  <physicalDescription>
    <internetMediaType>text/html</internetMediaType>
    <internetMediaType>image/jpg</internetMediaType>
  </physicalDescription>
  <abstract>Web site promoting the candidacy of Fran Ulmer, Democratic candidate for Governor, Alaska, 2002.
Includes candidate biography, issue position statements, campaign contact information, privacy policy and
campaign news press releases. Site features enable visitors to sign up for campaign email list, volunteer, make
campaign contributions and follow links to other internet locations. </abstract>
  <subject>
    <topic>Elections</topic>
    <geographic>Alaska</geographic>
  </subject>
  <subject>
    <topic>Governors</topic>
    <geographic>Alaska</geographic>
    <topic>Election</topic>
  </subject>
  <subject>
    <topic>Democratic Party (AK)</topic>
  </subject>
  <relatedItem type="host">
    <titleInfo>
      <title>Election 2002 Web Archive </title>
    </titleInfo>
    <identifier type="uri"> http://www.loc.gov/minerva/collect/elec2002/ </identifier>
  </relatedItem>
  <identifier type="uri" displayLabel="Active site (if available)">http://www.franulmer.com/ </identifier>
  <identifier type="uri" displayLabel="Archived site">http://wayback-
cgi1.alexa.com/e2002/*/http://www.franulmer.com/ </identifier>
  <accessCondition/>
</mods>

```

Web archive (Figure 1), the user display of the same record (Figure 2), and the result of following the link to the archived Web sites (Figure 3).

Metadata for additional collections

Although the 11 September Web archive was collected between 11 September, 2001 and 1 December, 2001 and the archived Web sites made available, the metadata is currently in the process of being completed and a search system established. This collection required different specifications for the content of fields,

particularly the title, subject headings, and languages. As with Election 2002, the contractor wanted to generate as much metadata as possible (for instance, the abstract and constructed title are generated from data concerning the producer and producer type for each site). The sites in the 11 September Web archive were much more diverse than the Election 2002 collection, so more individual analysis of sites was necessary and the metadata is not as rich. For instance, it is only possible to supply very general subject headings for each individual site without detailed subject analysis.

Figure 2 Record display for Fran Ulmer Web site

Election 2002 Web Archive Record

Title: [FranUlmer.com -- Home Page](#)

Alternative Title: [Fran Ulmer, Democratic candidate for Governor, Alaska, 2002](#)

Name: **Ulmer, Fran**

Abstract: Web site promoting the candidacy of Fran Ulmer, Democratic candidate for Governor, Alaska, 2002. Includes candidate biography, issue position statements, campaign contact information, privacy policy and campaign news press releases. Site features enable visitors to sign up for campaign email list, volunteer, make campaign contributions and follow links to other internet locations.

Date Captured: July 2, 2002 - December 3, 2002
[Archived Site](#)

Subjects: Elections--Alaska
Governors--Alaska--Election
Democratic Party (AK)

Language: English

Genre: Web site

Access Condition: None

Active Site: <http://www.franulmer.com/>

Collection Title: [Election 2002 Web Archive](#)

Figure 3 Resource page for archived Web site

Election 2002 Web Archive Resource Page

Each date listed below represents one archived version of « <http://www.franulmer.com/> ». Clicking any date will display the web site record as it appeared at that time. Following other links from that page will return the closest matches available from the Election 2002 Web Archive. The record will open in the same window. Please use the browser back button to return to this page if you wish to visit other archived dates.

194 archived captures for <http://www.franulmer.com/>

* denotes when site was updated.

Search Results for 2002					
Jan-Feb	Mar-Apr	May-Jun	Jul-Aug	Sep-Oct	Nov-Dec
0 pages	0 pages	0 pages	57 pages	62 pages	75 pages
			Jul 02, 2002 *	Sep 01, 2002	Nov 01, 2002 *
			Jul 03, 2002 *	Sep 02, 2002	Nov 02, 2002 *
			Jul 03, 2002	Sep 03, 2002	Nov 03, 2002
			Jul 04, 2002	Sep 04, 2002	Nov 04, 2002 *
			Jul 05, 2002	Sep 04, 2002	Nov 05, 2002
			Jul 07, 2002	Sep 05, 2002 *	Nov 05, 2002
			Jul 09, 2002	Sep 06, 2002 *	Nov 05, 2002
			Jul 10, 2002	Sep 07, 2002	Nov 05, 2002
			Jul 11, 2002	Sep 08, 2002	Nov 05, 2002 *
			Jul 12, 2002 *	Sep 09, 2002	Nov 05, 2002

The Library of Congress is planning to provide metadata using MODS for the 107th Congress Web archive by using tools developed in-house. The resulting records will be of high quality because LC staff will be providing the content and will apply controlled vocabulary (name and subject headings). METS objects will be created for the sites to include technical

and administrative metadata as well as descriptive.

Conclusions

MODS draws on the well-developed bibliographic description traditions that the

library community has contributed to the organization of knowledge. By retaining some of the richness of the MARC element set and replacing the syntax with XML and the more friendly language-based tags, MODS allows for rich resource description that is compatible with the huge numbers of MARC bibliographic records existing in library systems. In addition, as an XML descriptive standard, it provides the flexibility to be combined with other XML-based standards such as METS to satisfy needs for the digital library environment.

Because of the detailed documentation that the Library of Congress provides in the MODS guidelines, mappings, and examples, MODS can be deployed by users without having to provide their own application instructions. In addition, tools are provided for transformations to and from MARC and to other metadata schemes. LC has a firm commitment to support, maintain and document its standards. As MODS is used more widely and for other applications within and outside LC, it will evolve to fill the need for a rich descriptive standard that is compatible with MARC and that benefits from the flexibility and power of the XML syntax.

References

- Ayres, M.-L. (2003), *Message to MODS List*, e-mail message, 25 August.
- Beaubien, R. (2003), *Message to MODS List*, e-mail message, 14 July.
- Dublin Core Metadata Initiative (2003), *Dublin Core Metadata Element Set, Version 1.1: Reference Description*, Web document, available at: www.dublincore.org/documents/dces (accessed 12 September 2003).
- Library of Congress (n.d., a), *Metadata Object Description Schema. MODS User Guidelines*, Web document, available at: www.loc.gov/standards/mods/mods-userguide.html (accessed 12 September 2003).
- Library of Congress (n.d., b), *Metadata Object Description Schema. MODS Types*, Web document, available at: www.loc.gov/standards/mods/mods-notes.html (accessed 12 September 2003).
- Library of Congress (n.d., c), *MINERVA: Mapping the Internet Electronic Resources Virtual Archive*, Web document, available at: www.loc.gov/minerval/ (accessed 12 September 2003).
- Library of Congress (n.d., d), *Digital Audio-Visual Preservation Prototyping Project*, Web document, available at: <http://lcweb.loc.gov/rr/mopic/avprot/> (accessed 12 September 2003).
- Library of Congress (n.d., e), *American Memory. Sample Requests Using the Open Archives Initiative Protocol for Metadata Harvesting*, Web document, available at: http://memory.loc.gov/ammem/oamh/oai_request.html (accessed 12 September 2003).
- Olson, T. (2002), *Message to MODS list*, e-mail message, 26 June.
- University of California at Berkeley (n.d.), *GDM: A Descriptive Metadata Extension Schema for the Metadata Encoding and Transmission Standard*, Web document, available at: <http://sunsite.berkeley.edu/mets/gdmSchema.html> (accessed 12 September 2003).