An introduction to the Metadata Encoding and Transmission Standard (METS)

Morgan V. Cundiff

The author

Morgan V. Cundiff is Senior Standards Specialist in the Network Development and MARC Standards Office (NDMSO), Library of Congress, Washington, DC, USA.

Keywords

Online cataloguing, Digital libraries, Standards

Abstract

This article provides an introductory overview of the Metadata Encoding and Transmission Standard, better known as METS. It will be of most use to librarians and technical staff who are encountering METS for the first time. The article contains a brief history of the development of METS, a primer covering the basic structure and content of METS documents, and a discussion of several issues relevant to the implementation and continuing development of METS including object models, extension schemata, and application profiles.

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

In the late 1990s several large university libraries (led by the University of California of Berkeley and with the support of the Digital Library Federation) worked together on a project known as Making of America II (MoA II). One of the objectives of MoA II was:

> . . . to create a proposed digital library object standard by encoding defined descriptive, administrative and structural metadata, along with primary content, inside a digital library object (University of California, 2001).

The project participants rightly articulated the need for such a standard encoding: to serve as a digital object transfer syntax, to function as a data format for use with digital libraries, and to function as a data format for use with digital repositories. To this end, MoA II produced an XML DTD that specified the data elements and encoding for a limited set of object types (only objects comprising text and still image content files, including diaries, journals, photographs, and correspondence). Several institutions undertook projects to create testbeds of digital objects based on the DTD, and also to develop software tools that supported the creation and processing of the object documents (Hurley *et al.*, 2000).

In February 2001 the Digital Library Federation sponsored a meeting called "The Making of America II DTD Workshop" (McDonough *et al.*, 2001). The meeting was planned and directed by Jerome McDonough of New York University for the purpose of considering the future of the MoA II DTD (McDonough, formerly of UC Berkeley, was the primary author of the MoA II DTD). The meeting was attended by both veterans of the Making of America II project as well as representatives of other institutions involved with digital library projects or with digital preservation. The question that the group considered was: Should the MoA II DTD be revised to address its identified shortcomings or should the MoA II DTD be deprecated in favor of one of the other emerging XML standards for digital object and/or metadata management such as MPEG-7, RDF, or SMIL? The outcome of the meeting was a consensus that

Library Hi Tech
Volume 22 · Number 1 · 2004 · pp. 52-64
Emerald Group Publishing Limited · ISSN 0737-8831
DOI 10.1108/07378830410524495

Received 17 September 2003
Revised 2 November 2003
Accepted 7 November 2003
© Morgan V. Cundiff.

Library Hi Tech
Volume 22 · Number 1 · 2004 · 52-64

the library community would be best served with a new version of the MoA II DTD. It was further decided that the DTD should be recast as an XML schema, and that the schema would not provide (as did MoA II) a vocabulary for expressing descriptive or administrative metadata. Rather, descriptive and administrative metadata would be (optionally) expressed using vocabularies specified by other appropriate schemata (either existing or yet to be created) and included in a given document instance by means of the XML namespace facility.

By April 2001 McDonough had produced a new draft schema for community review and it had been decided to call it Metadata Encoding and Transmission Standard (METS) (Library of Congress, 2003a). The ongoing development of METS continues to be an initiative of the Digital Library Federation. A METS editorial board was established with McDonough serving as chair. It was also arranged for the Library of Congress Network Development and MARC Standards Office (NDMSO) to be the official maintenance agency for the standard. The Library of Congress maintains the official METS Web site (www.loc.gov/standards/mets/) and listserv. Development work on METS has been continuous to the present. Development has been an open process with comments and suggestions coming from many different contributors. The current version of METS is 1.3 and is considered to be stable (Library of Congress, n.d.).

A METS primer

A definition of METS might go something like this:

> METS is an XML schema designed for the purpose of creating XML document instances that express the hierarchical structure of digital library objects, the names and locations of the files that comprise those digital objects, and the associated descriptive and administrative metadata. (Note that, at present, METS is limited to digital objects comprising text, images, audio, and video files.)

This definition immediately begs the question, "What is a digital library object?" This is a question that could lead to a very open-ended discussion. METS itself does not provide a strict definition. However, for the purpose of the examples that follow, a digital library object (and therefore a METS document as well) will have a one-to-one correspondence with a typical library item (e.g. a book, a photograph, a sound recording, a map, and so on). This is in fact the approach most often taken with the several METS projects to date. More will be said below about other possibilities.

Sections of a METS document

A METS document is most quickly understood by examining its basic structural components[1]. A METS document can have up to seven major subsections:

(1) a Mets Header (metsHdr);
(2) a Descriptive Metadata Section (dmdSec);
(3) an Administrative Metadata Section (amdSec);
(4) a File Section (fileSec);
(5) a Structure Map (structMap);
(6) Structural Links (structLink); and
(7) a Behavior Section (behaviorSec) (see Figure 1).

The Structure Map

At a minimum, a METS document must have one major subsection, a Structure Map (structMap). The Structure Map is the backbone of a METS document. It is the means by which the hierarchical structure and the sequence of the components of a digital object are expressed. This is achieved by use of nested division (div) elements. Typical structural divisions of a digital library object might be the chapters and pages that comprise a book or the separate tracks that comprise a compact disc. A simple sketch (Figure 2) illustrates this.

Figure 3 shows a valid METS document for a compact disc with two tracks. Notice that the TYPE attribute indicates the type of div and the

Figure 1 Sketch of METS with seven major subsections

```
<mets>
  <metsHdr/>
  <dmdSec/>
  <amdSec/>
  <fileSec/>
  <structMap/>
  <structLink/>
  <behaviorSec/>
</mets>
```

Library Hi Tech
Volume 22 · Number 1 · 2004 · 52-64

Figure 2 Sketch of minimal METS document with Structure Map (structMap)

```
<mets>
  <structMap>
    <div>
      <div></div>
    </div>
  </structMap>
</mets>
```

Figure 3 A valid METS document with Structure Map (structMap)

```
<?xml version="1.0" encoding="UTF-8"?>
<mets:mets
  xmlns:mets="http://www.loc.gov/METS/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:xlink="http://www.w3.org/TR/xlink"
  xsi:schemaLocation="http://www.loc.gov/METS/
  http://www.loc.gov/standards/mets/mets.xsd
  ">
  <mets:structMap>
    <mets:div TYPE="compactDiscObject">
      <mets:div ORDER="01" TYPE="track"/>
      <mets:div ORDER="02" TYPE="track"/>
    </mets:div>
  </mets:structMap>
</mets:mets>
```

ORDER attribute indicates the sequence in which the tracks on the compact discs occur.

The File Section

The document in Figure 3 is a valid METS instance but it is not yet very useful. Though it is not required, a METS object will most often also include a File Section (fileSec) element, as illustrated in Figure 4.

The File Section is where the names and locations of the files that comprise the digital object are listed. Figure 5 shows the compact disc example with a File Group (fileGrp) added.

Notice that the fileSec element contains one File Group (fileGrp) element and that the fileGrp element contains two File (file) elements, one for each of the two tracks on the compact disc. Each file element contains a File Location (Flocat) element, which specifies the location and filename of a file. Notice also in the structMap that each of the two div elements for the compact disc tracks contains a File Pointer (fptr) element. Within a METS document it is often desirable to associate information in one section of the document with information in a different section. This internal linking is accomplished by means of XML ID/IDREF attributes. In this case, each of the fptr elements has a FILEID attribute (an IDREF attribute), which contains a value (e.g. "file01") that makes the association between the particular div and the corresponding file. Here, the corresponding file element has an ID attribute with the matching value "file01". The METS document is now much more useful. It is now possible to discern the structure of the digital object and also to locate the corresponding files[2].

Figure 4 Sketch of METS document with File Section (fileSec) element

```
<mets>
  <fileSec></fileSec>
  <structMap></structMap>
</mets>
```

Figure 5 METS document with File Section (fileSec) element

```
<?xml version="1.0" encoding="UTF-8"?>
<mets:mets
  xmlns:mets="http://www.loc.gov/METS/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:xlink="http://www.w3.org/TR/xlink"
  xsi:schemaLocation="
  http://www.w3.org/TR/xlink  http://www.loc.gov/standards/mets/xlink.xsd
  http://www.loc.gov/METS/  http://www.loc.gov/standards/mets/mets.xsd
  ">
  <mets:fileSec>
    <mets:fileGrp>
      <mets:file ID="file01">
        <mets:FLocat LOCTYPE="URL" xlink:href="http://path/001.wav"/>
      </mets:file>
      <mets:file ID="file02">
        <mets:FLocat LOCTYPE="URL" xlink:href="http://path/002.wav"/>
      </mets:file>
    </mets:fileGrp>
  </mets:fileSec>
  <mets:structMap>
    <mets:div TYPE="compactDiscObject">
      <mets:div ORDER="01" TYPE="track">
        <mets:fptr FILEID="file01"/>
      </mets:div>
      <mets:div ORDER="02" TYPE="track">
        <mets:fptr FILEID="file02"/>
      </mets:div>
    </mets:div>
  </mets:structMap>
</mets:mets>
```

Descriptive Metadata

A METS document may optionally include a Descriptive Metadata Section (dmdSec), as illustrated in Figure 6.

Library Hi Tech
Volume 22 · Number 1 · 2004 · 52-64

Figure 6 Sketch of METS document with Descriptive Metadata Section (dmdSec) element

```
<mets>
  <dmdSec></dmdSec>
  <fileSec></fileSec>
  <structMap></structMap>
</mets>
```

The Descriptive Metadata Section does not provide a descriptive metadata vocabulary of its own. Rather, it provides a mechanism for pointing to metadata in external documents or systems using the Metadata Reference (mdRef) element, and also, alternatively, provides a mechanism for embedding descriptive metadata from a different namespace in the METS document using the Metadata Wrap (mdWrap) element. The sketch in Figure 7 shows the use of the mdRef element.

Figure 8 shows a METS document with a Metadata Reference (mdRef) element. Notice that the xlink:href attribute of the mdRef element contains a URL which points to a MARC bibliographic record in an external bibliographic database.

An alternative method of managing descriptive metadata is to embed the bibliographic data into the METS document. This is accomplished by using the Metadata Wrap (mdWrap) element and either the XML Data (xmlData) or Binary Data (binData) sub-elements. The former is probably more typical and is illustrated here (Figure 9). The XML Data (xmlData) element acts as a kind of "socket" into which elements from a different namespace (i.e. elements defined in a different XML schema) can be inserted.

Figure 10 shows a METS document with elements from the MODS descriptive metadata schema. In this case MODS is functioning as an "extension schema" to METS (more will be said about extension schemata below).

Figure 7 Sketch of METS document with Descriptive Metadata Section (dmdSec) element

```
<mets>
  <dmdSec>
    <mdRef/>
  </dmdSec>
  <fileSec></fileSec>
  <structMap></structMap>
</mets>
```

Notice that the elements from the MODS schema (titleInfo, physicalDescription, etc.) are "namespace qualified" with the "mods:" prefix, indicating that these elements belong to the MODS vocabulary and not to METS. Notice also the pairs of ID/IDREF attributes. The fptr FILEID points to the file ID and the div DMDID points to the dmdSec ID. Making these associations explicit in this example may not really be necessary, because there is only one div, one file, and one dmdSec. However, in cases where there are multiple div elements, file elements, and perhaps multiple sections of descriptive metadata, the associations are more important. Consider the example in Figure 10 in which there are two div elements with fptr subelements, two file elements, and two MODS relatedItem elements containing descriptive metadata for each of the two tracks on the compact disc. In a case like this ID/IDREF links are necessary for maintaining the appropriate associations between different metadata sections.

Notice the ID/IDREF links from fptr to file elements (as in Figure 9) but also notice the links from the div elements to dmdSec elements. The high-level div (TYPE= "compactDiscObject") is associated with the entire dmdSec and the lower-level div elements (TYPE="track") are associated with track-level descriptive metadata. This ability to make associations between different levels of the Structure Map with the appropriate sub-parts of the descriptive metadata is a very powerful and useful feature (see Figure 11).

Administrative metadata

As with descriptive metadata, METS does not provide a vocabulary of its own for administrative metadata. The Administrative Metadata Section (amdSec)is further divided into four subsections:

(1) Technical Metadata (techMD);
(2) Rights Metadata (rightsMD);
(3) Source Metadata (sourceMD); and
(4) Digital Provenance Metadata (digiprovMD).

Again, as with descriptive metadata, each of the four administrative metadata subsections may contain a Metadata Reference (mdRef) subelement to point to metadata external to the

Library Hi Tech
Volume 22 · Number 1 · 2004 · 52-64

Figure 8 METS document with Metadata Reference (mdRef) element (points to descriptive metadata record in external system)

```
<?xml version="1.0" encoding="UTF-8"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xlink="http://www.w3.org/TR/xlink" xsi:schemaLocation="
  http://www.w3.org/TR/xlink  http://www.loc.gov/standards/mets/xlink.xsd
  http://www.loc.gov/METS/  http://www.loc.gov/standards/mets/mets.xsd
  ">
  <mets:dmdSec ID="dmd001">
    <mets:mdRef LOCTYPE="URN" MDTYPE="MARC"
xlink:href="http://lcweb2.loc.gov/cgi-
bin/query/r?pp/ils:@field(NUMBER+@band(cph+3c17261))"/>
  </mets:dmdSec>
  <mets:fileSec>
    <mets:fileGrp>
      <mets:file ID="file01">
        <mets:FLocat LOCTYPE="URL"
          xlink:href="http://memory.loc.gov/service/pnp/cph/3c10000/3c17000/3c17200/3c17261v.jpg"/>
      </mets:file>
    </mets:fileGrp>
  </mets:fileSec>
  <mets:structMap>
    <mets:div TYPE="photoObject">
      <mets:fptr FILEID="file01"/>
    </mets:div>
  </mets:structMap>
</mets:mets>
```

Figure 9 Sketch of METS document with XML Data (xmlData) element

```
<mets>
  <dmdSec>
    <mdWrap>
      <xmlData>
       <!-- insert data from different namespace here -->
      </xmlData>
    </mdWrap>
  </dmdSec>
  <fileSec></fileSec>
  <structMap></structMap>
</mets>
```

METS document, or a Metadata Wrap (mdWrap) sub-element to embed administrative metadata in the METS document. And again, it is with the use of extension schemata that administrative metadata are expressed within a METS document (see sketch in Figure 12).

Figure 13 shows a METS document for a single digital still image (as in Figure 10). This example uses the MIX schema to express technical metadata for still images. Including rights metadata, source metadata, digital provenance metadata would all work the same way, though of course using different extension schemata. Notice that the file element ADMID attribute value ("tmd001") associates the file with the appropriate technical metadata.

METS Header

A METS document may optionally include a Mets Header (metsHdr) section. The METS Header contains information about the METS document itself (primarily information about the creation of the document) as opposed to information about the source object (see Figure 14).

Structural Links

A METS document may optionally include a Structural Links section. The Structural Links (structLink) element may contain one or more Structural Map Link (smLink) elements, each of which can express a link between any two div elements in the Structure Map (structMap) section. This facility was added to METS to provide a way to express the hyperlinks between Web pages in the case where the digital object in question is an archived Web site, or "Web site snapshot". A simple example of this is provided in Figure 14. It could possibly be used

Library Hi Tech
Volume 22 · Number 1 · 2004 · 52-64

Figure 10 METS document with Metadata Wrap (mdWrap) element

```
<?xml version="1.0" encoding="UTF-8"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/" xmlns:mods="http://www.loc.gov/mods/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xlink="http://www.w3.org/TR/xlink" xsi:schemaLocation="
 http://www.w3.org/TR/xlink  http://www.loc.gov/standards/mets/xlink.xsd
 http://www.loc.gov/mods/  http://www.loc.gov/standards/mods/mods.xsd
 http://www.loc.gov/METS/  http://www.loc.gov/standards/mets/mets.xsd
 http://www.w3.org/XML/1998/namespace  http://www.w3.org/2001/xml.xsd
 ">
 <mets:dmdSec ID="dmd001">
  <mets:mdWrap MDTYPE="MARC">
   <mets:xmlData>
    <mods:mods>
     <mods:titleInfo>
      <mods:title>Like modern town criers, policemen boom the news
        to depositors that bank is closed [graphic]</mods:title>
     </mods:titleInfo>
     <mods:physicalDescription>
      <mods:extent>1 photograph.</mods:extent>
     </mods:physicalDescription>
    </mods:mods>
   </mets:xmlData>
  </mets:mdWrap>
 </mets:dmdSec>
 <mets:fileSec>
  <mets:fileGrp>
   <mets:file ID="file001">
    <mets:FLocat LOCTYPE="URL" xlink:href="c:\alchemy\images\3c30861u.tif"/>
   </mets:file>
  </mets:fileGrp>
 </mets:fileSec>
 <mets:structMap>
  <mets:div DMDID="dmd001">
   <mets:fptr FILEID="file001"/>
  </mets:div>
 </mets:structMap>
</mets:mets>
```

to create "alternative, horizontal arrangements of the divisions of the structMap" (Beaubien, 2003) for any object type, if there was a reason to do so (see Figure 15).

Behavior Section

A METS document may optionally include a Behavior Section (behaviorSec):

> A behavior section can be used to associate executable behaviors with content in the METS object. A behavior section contains one or more behavior elements, each of which has an interface definition element that represents an abstract definition of the set of behaviors represented by a particular behavior section. A behavior also has a mechanism element which is used to point to a module of executable code that implements and runs the behavior defined abstractly by the interface definition. Digital object behaviors can be implemented as linkages to distributed Web services (Library of Congress, 2003b) (see Figure 15).

The Behavior Section was added to METS at the urging of the FEDORA project. The FEDORA Web site provides additional information about and examples of the Behavior Section (www.fedora.info). An example of a behavior might be something like a page-turning functionality for a book object (see Figure 16).

Notice that the behavior element STRUCTID attribute is used to associate the behavior with a particular div element.

Object models

In the examples given thus far there has been a simple relationship between one library item (a

Figure 11 METS document with Metadata Wrap (mdWrap) element

```
<?xml version="1.0" encoding="UTF-8"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/"
 xmlns:mods="http://www.loc.gov/mods/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:xlink="http://www.w3.org/TR/xlink" xsi:schemaLocation="
  http://www.w3.org/TR/xlink  http://www.loc.gov/standards/mets/xlink.xsd
  http://www.loc.gov/mods/  http://www.loc.gov/standards/mods/mods.xsd
  http://www.loc.gov/METS/  http://www.loc.gov/standards/mets/mets.xsd
  http://www.w3.org/XML/1998/namespace  http://www.w3.org/2001/xml.xsd
  ">
  <mets:dmdSec ID="dmd001">
    <mets:mdWrap MDTYPE="MODS">
      <mets:xmlData>
        <mods:mods>
          <mods:titleInfo>
            <mods:title>The Best of Nimbus Records 1993</mods:title>
          </mods:titleInfo>
          <mods:relatedItem ID="tr001">
            <mods:titleInfo>
              <mods:title>Eclogue for piano and strings</mods:title>
            </mods:titleInfo>
          </mods:relatedItem>
          <mods:relatedItem ID="tr002">
            <mods:titleInfo>
              <mods:title>Jubilte Deo</mods:title>
            </mods:titleInfo>
          </mods:relatedItem>
        </mods:mods>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:dmdSec>
  <mets:fileSec>
    <mets:fileGrp>
      <mets:file ID="file01">
        <mets:FLocat LOCTYPE="URL" xlink:href="http://www.loc.gov/path/001.wav"/>
      </mets:file>
      <mets:file ID="file02">
        <mets:FLocat LOCTYPE="URL" xlink:href="http://www.loc.gov/path/002.wav"/>
      </mets:file>
    </mets:fileGrp>
  </mets:fileSec>
  <mets:structMap>
    <mets:div TYPE="compactDiscObject" DMDID="dmd001">
      <mets:div ORDER="01" TYPE="track" DMDID="tr001">
        <mets:fptr FILEID="file01"/>
      </mets:div>
      <mets:div ORDER="02" TYPE="track" DMDID="tr002">
        <mets:fptr FILEID="file02"/>
      </mets:div>
    </mets:div>
  </mets:structMap>
</mets:mets>
```

compact disc, a photograph, etc.) and one digital library object, i.e. one library item is expressed as one METS document. However, different approaches are possible. In the examples above the fptr element was used to associate a particular div element with a particular file element within a single METS document. It is also possible to make associations between div elements and other METS documents using the METS Pointer (mptr) element. This makes it possible to create parent-child relationships and thus to create hierarchies of METS documents. There are several scenarios where this approach might be

Library Hi Tech
Volume 22 · Number 1 · 2004 · 52-64

Figure 12 Sketch of METS document with techMD, rightsMD, sourceMD, and digiprovMD elements

```
<mets>
 <amdSec>
  <techMD>
    <mdWrap>
      <xmlData>
        <!-- insert data from different namespace here -->
      </xmlData>
    </mdWrap>
  </techMD>
 </amdSec>
 <amdSec>
  <rightsMD>
    <mdWrap>
      <xmlData>
        <!-- insert data from different namespace here -->
      </xmlData>
    </mdWrap>
  </rightsMD>
 </amdSec>
 <amdSec>
  <sourceMD>
    <mdWrap>
      <xmlData>
        <!-- insert data from different namespace here -->
      </xmlData>
    </mdWrap>
  </sourceMD>
 </amdSec>
 <amdSec>
  <digiprovMD>
    <mdWrap>
      <xmlData>
        <!-- insert data from different namespace here -->
      </xmlData>
    </mdWrap>
  </digiprovMD>
 </amdSec>
 <fileSec></fileSec>
 <structMap></structMap>
</mets>
```

used. One is to encode a journal run using a top-down hierarchy of Title to Volumes to Issues to Articles. Each level of the hierarchy would comprise one or more METS documents. Another scenario might be where, in a particular repository, there is reason to create a METS document for every file and then to associate those files at the object level with a parent METS document. Another scenario is use a METS document to aggregate any arbitrary group of digital objects into a collection. At the Library of Congress this approach was taken for a recent project called "Patriotic Melodies" (www.loc.gov/rr/perform/patriotic). Collection objects were created for a given work and then library item objects were

Figure 13 METS document with Technical Metadata (techMD) element

```
<?xml version="1.0" encoding="UTF-8"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/"
 xmlns:mods="http://www.loc.gov/mods/"
 xmlns:mix="http://www.loc.gov/mix/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:xlink="http://www.w3.org/TR/xlink" xsi:schemaLocation="
 http://www.w3.org/TR/xlink  http://www.loc.gov/standards/mets/xlink.xsd
 http://www.loc.gov/mods/  http://www.loc.gov/standards/mods/mods.xsd
 http://www.loc.gov/METS/  http://www.loc.gov/standards/mets/mets.xsd
 http://www.loc.gov/mix/  http://www.loc.gov/standards/mix/mix.xsd
 http://www.w3.org/XML/1998/namespace  http://www.w3.org/2001/xml.xsd
 ">
 <mets:dmdSec ID="dmd001">
  <mets:mdWrap MDTYPE="MODS">
   <mets:xmlData>
    <mods:mods>
     <mods:titleInfo>
      <mods:title>Like modern town criers, policemen boom the news
       to depositors that bank is closed [graphic]</mods:title>
     </mods:titleInfo>
     <mods:physicalDescription>
      <mods:extent>1 photograph.</mods:extent>
     </mods:physicalDescription>
    </mods:mods>
   </mets:xmlData>
  </mets:mdWrap>
 </mets:dmdSec>
 <mets:amdSec>
  <mets:techMD ID="tmd001">
   <mets:mdWrap MDTYPE="NISOIMG">
    <mets:xmlData>
     <mix:mix>
      <mix:BasicImageParameters>
       <mix:Format>
        <mix:MIMEType>image/tiff</mix:MIMEType>
        <mix:ByteOrder>big-endian</mix:ByteOrder>
        <mix:Compression>
         <mix:CompressionScheme>32773</mix:CompressionScheme>
        </mix:Compression>
        <mix:PhotometricInterpretation>
         <mix:ColorSpace>2</mix:ColorSpace>
        </mix:PhotometricInterpretation>
        <mix:Segments>
         <mix:SegmentType>0</mix:SegmentType>
         <mix:StripOffsets>246</mix:StripOffsets>
         <mix:RowsPerStrip>16</mix:RowsPerStrip>
         <mix:StripByteCounts>1242</mix:StripByteCounts>
        </mix:Segments>
       </mix:Format>
       <mix:File>
        <mix:FileSize>5932</mix:FileSize>
        <mix:Checksum>
         <mix:ChecksumMethod>MD5</mix:ChecksumMethod>
         <mix:ChecksumValue>37bfffadd6087139139f0196c775dd0b
         </mix:ChecksumValue>
        </mix:Checksum>
       </mix:File>
      </mix:BasicImageParameters>
     </mix:mix>
    </mets:xmlData>
   </mets:mdWrap>
  </mets:techMD>
 </mets:amdSec>
 <mets:fileSec>
  <mets:fileGrp>
   <mets:file ID="file001" ADMID="tmd001">
    <mets:FLocat LOCTYPE="URL" xlink:href="c:\alchemy\images\3c30861u.tif"/>
   </mets:file>
  </mets:fileGrp>
 </mets:fileSec>
 <mets:structMap>
  <mets:div DMDID="dmd001">
   <mets:fptr FILEID="file001"/>
  </mets:div>
 </mets:structMap>
</mets:mets>
```

Figure 14 METS document with METS Header (metsHdr) element

```
<?xml version="1.0" encoding="utf-8"?>
<mets:mets
  xmlns:mets="http://www.loc.gov/METS/"
  xmlns:xlink="http://www.w3.org/TR/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="
  http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd
  ">
  <mets:metsHdr CREATEDATE="2002-09-12T13:20:00-05:00">
    <mets:agent ROLE="CREATOR" TYPE="ORGANIZATION">
      <mets:name>Library of Congress</mets:name>
    </mets:agent>
  </mets:metsHdr>
  <mets:fileSec>
    <mets:fileGrp/>
  </mets:fileSec>
  <mets:structMap>
    <mets:div/>
  </mets:structMap>
</mets:mets>
```

Figure 15 METS document with Structural Links (structLink) element

```
<?xml version="1.0" encoding="utf-8"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/"
  xmlns:xlink="http://www.w3.org/TR/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="
  http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd
  ">
  <mets:fileSec>
    <mets:fileGrp/>
  </mets:fileSec>
  <mets:structMap>
    <mets:div ID="div001" TYPE="home_page">
      <mets:div ID="div002a" TYPE="linked_page"/>
      <mets:div ID="div002b" TYPE="linked_page"/>
      <mets:div ID="div002c" TYPE="linked_page"/>
      <mets:div ID="div002d" TYPE="linked_page"/>
    </mets:div>
  </mets:structMap>
  <mets:structLink>
    <mets:smLink from="div001" to="div002a" xlink:title="Home to linked page"/>
    <mets:smLink from="div001" to="div002b" xlink:title="Home to linked page"/>
    <mets:smLink from="div001" to="div002c" xlink:title="Home to linked page"/>
    <mets:smLink from="div001" to="div002d" xlink:title="Home to linked page"/>
    <mets:smLink from="div002a" to="div001" xlink:title="Return to home page"/>
    <mets:smLink from="div002b" to="div001" xlink:title="Return to home page"/>
    <mets:smLink from="div002c" to="div001" xlink:title="Return to home page"/>
    <mets:smLink from="div002d" to="div001" xlink:title="Return to home page"/>
    <mets:smLink from="div002a" to="div002b" xlink:title="Next"/>
    <mets:smLink from="div002b" to="div002c" xlink:title="Next"/>
    <mets:smLink from="div002c" to="div002d" xlink:title="Next"/>
    <mets:smLink from="div002d" to="div002a" xlink:title="Next"/>
    <mets:smLink from="div002a" to="div002d" xlink:title="Previous"/>
    <mets:smLink from="div002d" to="div002c" xlink:title="Previous"/>
    <mets:smLink from="div002c" to="div002d" xlink:title="Previous"/>
    <mets:smLink from="div002b" to="div002a" xlink:title="Previous"/>
  </mets:structLink>
</mets:mets>
```

Figure 16 METS document with Behavior Section (behaviorSec) element

```
<?xml version="1.0" encoding="utf-8"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/"
  xmlns:xlink="http://www.w3.org/TR/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="
  http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd
  ">
  <mets:fileSec>
    <mets:fileGrp/>
  </mets:fileSec>
  <mets:structMap>
    <mets:div ID="div001"/>
  </mets:structMap>
  <mets:behaviorSec>
    <mets:behavior ID="beh001" STRUCTID="div001" BTYPE="LC:simplePageTurner">
      <mets:interfaceDef LOCTYPE="URN" LABEL="LC:simplePageTurner Behavior Definition"
        xlink:href="LC:bdef:simplePageTurner"/>
      <mets:mechanism LOCTYPE="URN" LABEL="LC:simplePageTurner Mechanism"
        xlink:href="LC:bmech:simplePageTurner"/>
    </mets:behavior>
  </mets:behaviorSec>
</mets:mets>
```

specified as members of the collection object. The example in Figure 17 shows a collection object for *The Library of Congress March*. The mptr element is used to point to the three members of the collection: a set of score and parts, a sound recording, and a set of manuscript sketches.

Extension schemata, OAIS, and preservation

It was noted that METS is extended by the use of extension schemata for descriptive metadata and administrative metadata. It was also noted that administrative metadata are subdivided into four subcategories (technical, rights, source, and digital provenance). Technical metadata schemata created to date have been specific to still images, text, audio, or video and the same might be true for Source metadata. So there is potential for a fairly long list of needed schemata. Though METS implementers are free to use any extension schemata, it is also clear that standardization will be beneficial for digital library and digital repository interoperability. To this end the METS editorial board endorses selected schemata for use with METS. To date, three extension schemata for descriptive metadata have been endorsed (MARCXML, MODS, and Simple Dublin Core) and two schemata for technical metadata (MIX for still images and TextMD for text)[3]. Schemata for expressing descriptive metadata are fairly stable but producing a complete suite of schemata for expressing administrative metadata remains one of the challenges for the METS community (Library of Congress, 2003c).

It is also interesting to note that the development of administrative metadata extension schemata is intertwined with the issue of "preservation metadata". The Reference Model for an Open Archival Information System (OAIS) (ISO, 2002) articulates a set of abstract information categories that are part of an Archival Information Package (AIP). These conceptual categories are called Packaging Information, Content Information (including Representation Information) and Preservation Description Information (including Reference Information, Context Information, Provenance Information, and Fixity Information). In a study done for the Library of Congress it was argued that these conceptual categories would map to a METS document that made use of a full set of extension schemata (Library of Congress, 2003c). Many, though not all, of the preservation metadata elements would reside in the Administrative Metadata extension schemata. Thus, a METS document could be considered to be an implementation of an OAIS Information Package. There is, of course, still

Library Hi Tech
Volume 22 · Number 1 · 2004 · 52-64

Figure 17 METS document functioning as a collection object (with METS Pointer elements)

```
<?xml version="1.0" encoding="utf-8"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/"
xmlns:p="http://www.loc.gov/mets/profiles/patriotism" xmlns:lc="http://www.loc.gov/mets/profiles/"
xmlns:xlink="http://www.w3.org/TR/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance"
  xmlns:mods="http://www.loc.gov/mods/"
  PROFILE="lc:patriotismSongCollection"
  OBJID="loc.natlib.ihas.200000027"
  xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd
http://www.loc.gov/mods/ http://www.loc.gov/standards/mods/mods.xsd">
  <mets:dmdSec ID="DMD">
    <mets:mdWrap OTHERMDTYPE="MODS" MDTYPE="OTHER" ID="MODS1">
      <mets:xmlData>
        <mods:mods>
          <mods:titleInfo>
            <mods:title>The Library of Congress March</mods:title>
          </mods:titleInfo>
        </mods:mods>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:dmdSec>
  <mets:fileSec>
    <mets:fileGrp>
      <mets:file ID="BLURB1">
        <mets:FLocat
xlink:href="http://lcweb2.loc.gov/natlib/ihas/patriotism/collection/200000027/200000027.xhtml"
LOCTYPE="URL"/>
      </mets:file>
    </mets:fileGrp>
  </mets:fileSec>
  <mets:structMap>
    <mets:div DMDID="MODS1" TYPE="p:songCollection">
      <mets:div TYPE="p:blurb">
        <mets:fptr FILEID="BLURB1"/>
      </mets:div>
      <mets:div TYPE="p:collectionMembers">
        <mets:div TYPE="p:scoreObjects">
          <mets:div LABEL="Score and band parts" TYPE="lc:score">
            <mets:mptr xlink:href="loc.natlib.ihas.100010802" LOCTYPE="HANDLE"/>
          </mets:div>
        </mets:div>
        <mets:div TYPE="p:soundRecordingObjects">
          <mets:div LABEL="U. S. Marine Band, 2003" TYPE="lc:simpleAudio">
            <mets:mptr xlink:href="loc.natlib.ihas.100010804" LOCTYPE="HANDLE"/>
          </mets:div>
        </mets:div>
        <mets:div TYPE="p:otherObjects">
          <mets:div LABEL="Original score fragments" TYPE="lc:manuscriptSketch">
            <mets:mptr xlink:href="loc.natlib.ihas.100010805" LOCTYPE="HANDLE"/>
          </mets:div>
        </mets:div>
      </mets:div>
    </mets:div>
  </mets:structMap>
</mets:mets>
```

much work to be done in the area of defining preservation metadata, and consequently the task of creating definitive extension schemata for use with METS is very difficult. The PREMIS (PREservation Metadata: Implementation Strategies) Working Group II is:

> . . . aimed at the development of recommendations and best practices for implementing preservation metadata.

Library Hi Tech
Volume 22 · Number 1 · 2004 · 52-64

The work of this group may be helpful in clarifying what preservation metadata elements need to be captured for a given object type (www.oclc.org/research/projects/pmwg/).

Application profiles

METS is intentionally designed to provide a great deal of flexibility. It is understood that different implementers, or communities of implementers, may have very different objectives and constraints, and that therefore local practices will vary widely. On the other hand, there are also compelling reasons to promote common practice, including interoperability of digital libraries, software tool creation, and easier entry for institutions just getting started with METS. The METS Board has recently proposed the creation and registration of METS Application Profiles. The METS Profile 1.0 Requirements document provides the following definition:

> METS Profiles are intended to describe a class of METS documents in sufficient detail to provide both document authors and programmers the guidance they require to create and process METS documents conforming with a particular profile. Institutions creating profiles for their METS objects may register those profiles with the Library of Congress Network Development and MARC Standards Office after they are approved by the METS Editorial Board; such profiles will be made publicly available, so that those wishing to produce METS documents in accordance with a particular institution's requirements may examine them (Library of Congress, 2003d).

An Application Profile contains a specified set of elements:

- A URI which has been assigned to the profile.
- A short title for the class of profiled METS documents.
- An abstract.
- A date and time specifying when the profile was created.
- Contact information.
- A date when the profile was registered with the Library of Congress.
- An indication of any other profiles which may be related to the current profile.
- An enumeration of all extension schemata.
- An enumeration of any rules of descriptions along with details of where those rules apply.
- An enumeration of any controlled vocabularies.
- A description of any structural requirements regarding the construction of the METS object itself.
- A detailed description of the allowable technical characteristics of content files or executable behaviors.
- A description of affiliated tools.

A METS Application Profile is expressed as an XML document itself. An XML schema has been created for this purpose (Library of Congress, 2003e).

Conclusion

METS provides a method for aggregating all the metadata relevant to a given digital library object. A well-constructed METS object is then relatively easy to manage. Because METS is expressed in XML, there is a wide range of software, often free, that can be used to create, store, display, transform, navigate, query, or publish METS objects. The current frontier for METS work includes additional extension schema development, application profile development, and software tool development. There is also need for thorough documentation and training material. A number of libraries and library organizations have already launched METS-based applications, including OCLC, RLG, the California Digital Library, the Library of Congress, the University of California at Berkeley, the University of Graz, Harvard, MIT, Stanford, and Oxford University (University of California, n.d.). The degree to which the use of METS continues to spread will be one of the interesting things for which to watch in 2004.

Notes

1 There are several prerequisites to gaining a good understanding of METS. Familiarity with the XML standard and the adjunct standards for XML schema, XLink, and XML Namespaces is especially important.

2 The relationship between fptr and file elements may not be as simple as in the examples given (see Beaubien, 2003).
3 A pioneering project at the Library of Congress, the Digital Audio-Visual Preservation Prototyping Project, has proposed a complete set of METS extension schemata, available at: http://lcweb.loc.gov/rr/mopic/avprot/metsmenu2.html

References

Beaubien, R. (2003), "BrainDumpRickBeaubien" (/wiki/BrainDumpRickBeaubien?action=fullsearch&value=BrainDumpRickBeaubien&literal=1&case=1), available at: http://raymondyee.net/wiki/BrainDumpRickBeaubien

Hurley, B., Price-Wilkin, J., Proffitt, M. and Besser, H. (2000), "The making of America II testbed project: a digital library service model", available at: www.clir.org/pubs/reports/pub87/contents.html

ISO (2002), *Reference Model for an Open Archival Information System (OAIS)*, ISO, Geneva, January, available at: wwwclassic.ccsds.org/documents/pdf/CCSDS-650.0-B-1.pdf

Library of Congress (n.d.), *METS Schema*, available at: www.loc.gov/standards/mets/mets.xsd

Library of Congress (2003a), *METS: An Overview and Tutorial*, available at: www.loc.gov/standards/mets/METSOverview.v2.html

Library of Congress (2003b), *METS Editorial Board Endorses Extension Schemata for Use with METS*, available at: www.loc.gov/standards/mets/news061503.html

Library of Congress (2003c), *Archival Information Package (AIP) Design Study*, available at: http://lcweb.loc.gov/rr/mopic/avprot/AIP-Study_v19.pdf

Library of Congress (2003d), *METS Profile 1.0 Requirements*, available at: www.loc.gov/standards/mets/profile_docs/METS.profile.requirements.rtf

Library of Congress (2003e), *METS Profile Schema Version 1.0*, available at: www.loc.gov/standards/mets/profile_docs/mets.profile.v1-0.xsd

McDonough, J., Myrick, L. and Stedfeld, E. (2001), "Report on the making of America II DTD Digital Library Federation Workshop", available at: www.diglib.org/standards/metssum.pdf

University of California (n.d.), *METS Implementation Registry*, available at http://sunsite.berkeley.edu/mets/registry

University of California (2001), *The Making of America II*, available at: http://sunsite.berkeley.edu/MOA2/