# METS: Standardized Encoding for Digital Library Objects

JEROME P. MCDONOUGH

*Graduate School of Library & Information Science*

*University of Illinois, Urbana-Champaign*

*501 E. Daniel Street*

*Champaign, IL 61820*

jmcdonou@uiuc.edu
(217) 244-5916
FAX: (217) 244-3302

Abstract – METS is an XML document format intended for the encoding of complex objects within digital libraries. It provides the means to record all of the descriptive, administrative, structural and behavioral metadata needed to manage and provide access to complex digital content. While it was designed to promote interoperability of digital content between digital library systems and contribute to the preservation of digital library materials, a variety of practical barriers to achieving these goals remain. However, many of these obstacles are shared by other communities of practice, such as the eLearning community working on the IMS content packaging standards and the MPEG-21 community, and the digital library community faces a unique opportunity at the moment to work closely with others to try to improve the interoperability of our content not only with our own repository systems, but those being used by others.

*Keywords: METS, XML, Structural Metadata, Digital Preservation, Interoperability*

## Introduction

In 1997, the Library at the University of California at Berkeley embarked on a project in concert with several other members of the Digital Library Federation (DLF) in an attempt to move the entire membership of the DLF closer to their shared vision of a national digital library. The project included a testbed phase which was to "provide a means for the DLF to investigate, refine, and recommend metadata elements and encodings used to discover, display, and navigate digital archival objects." [7, p. 3] The testbed phase of this project resulted in the creation of the Making of America II (MOA2) DTD, an XML encoding format which recorded descriptive, administrative and structural metadata for the digital archival objects produced for the testbed.

While the MOA2 DTD sufficed for encoding the digital objects produced for the Making of America II project, it was limited in several ways. It provided no flexibility in terms of the exact metadata elements to be used for descriptive, administrative and structural metadata. Its facilities for structural metadata were also intentionally limited in scope to support for text and still image materials (whether single image or multiple image 'page-turned' objects); there was no attempt to support time-based media such as audio or video materials. These deficiencies led to a meeting of representatives from a variety of institutions pursuing digital library development programs at New York University in February of 2001 to investigate whether a successor format to the MOA2 DTD should be created. There was general agreement that such a path should be pursued, and this led to the creation of the Metadata Encoding and Transmission Standard (METS).

METS, like the MOA2 DTD, was designed to address several fairly specific needs of the digital library community. It needed to provide a framework for descriptive, administrative and structural metadata. It also needed to provide some flexibility for local practice with respect to descriptive and administrative metadata, while promoting standardization where possible. It had to provide mechanisms for encoding structural metadata for time-based media, in addition to support for text and still image materials. More generally, there was a desire for METS to facilitate the exchange and interoperability of digital library objects across digital library systems and to provide support for the long-term preservation of digital library objects by serving as a potential Submission Information Package, Archival Information Package and Dissemination Information Package within the Open Archival Information System Reference Model. [5]

At the same time that the digital library community was beginning to develop standards for the encoding of complex digital library objects to be used in a networked environment, a variety of other communities were engaged in similar endeavors. In 1997, when the MOA2 project was initiated, the National Learning Infrastructure Initiative of EDUCAUSE spun off a project called IMS to create specifications to support the use of digital content in distributed learning. Also in 1997, the United States' Office of the Secretary of Defense launched the Advanced Distributed Learning initiative to bring together government, industry and academia to promote the interoperability of learning tools and course content. These two initiatives led to the development of the IMS Global Learning Consortium specifications (including the IMS Content Packaging Specification [8]) and the Sharable Content Object Reference Model (SCORM)[1] respectively. In 1999, Subcommittee 29 of the ISO/IEC Joint Technical Committee 1 approved a new work item based on previous work on a 'multimedia framework.' This work item led to the creation of the MPEG-21 standard, including the MPEG-21 Digital Item Declaration Language.[10] The latter half of the 1990's saw a sudden flowering of standards for the encoding of complex digital objects.

This is not particularly surprising. With the spread of the Internet, numerous communities suddenly confronted both the opportunity to exchange information on a scale never before realized, and the necessity for establishing standards for encoding content in order to exchange information more sophisticated than an HTML page. The METS standard can be considered one of many efforts to try to determine, for one particular community, how complex sets of data and metadata might best be encoded to support both information exchange and information longevity. The remainder of this article provides an overview of the METS format and discusses some of the more critical problems that METS implementors may confront, including problems with

---

[1] See http://www.adlnet.org/index.cfm?fuseaction=scormabt for the complete set of documents comprising the SCORM specification.

interoperability that affect not only the METS standard but similar object encoding standards such as MPEG-21 and IMS Content Packaging, and problems in dealing with digital preservation metadata and complex object encoding.

## The METS Format

METS is an XML format defined by a schema expressed in the World Wide Web Consortium's XML Schema language. [16][2] The document format defined by the schema consists of seven major sections:

- A header, for metadata about the METS document itself;
- A descriptive metadata section;
- An administrative metadata section;
- An inventory of content files comprising the digital object;
- A structural linking section;
- A structural map; and
- A behavioral metadata section.

The administrative metadata section is further subdivided into sections for technical metadata regarding content files, intellectual property rights and permissions metadata, metadata regarding analog source material from which digital content has been derived, and a section for digital provenance metadata, which records life-cycle information regarding digital content.

The structural map is the core of a METS document, and is the only mandatory section in a METS file. The METS structural map owes a debt to the Text Encoding Initiative [15] default text structure and similarly relies on a series of nested division elements to record an abstract hierarchical structure for any given work. Lewis Carroll's *Hunting of the Snark: an Agony in Eight Fits* [3] might have a structural map as follows:

```
<structMap>
   <div ORDER="1" TYPE="Agony" LABEL="Hunting of the Snark">
      <div ORDER="1" TYPE="Fit" LABEL="The Landing" />
      <div ORDER="2" TYPE="Fit" LABEL="The Bellman's Speech" />
      <div ORDER="3" TYPE="Fit" LABEL="The Baker's Tale" />
      <div ORDER="4" TYPE="Fit" LABEL="The Hunting" />
      <div ORDER="5" TYPE="Fit" LABEL="The Beaver's Lesson" />
```

```
        <div ORDER="6" TYPE="Fit" LABEL="The Barrister's Dream" />
        <div ORDER="7" TYPE="Fit" LABEL="The Banker's Fate" />
        <div ORDER="8" TYPE="Fit" LABEL="The Vanishing" />
    </div>
</structMap>
```

A movie might be broken down into a similar hierarchical decomposition, with a single movie being comprised of multiple sequences, each consisting of one or more scenes, and with each scene consisting of one or more shots.[2]

This abstract structural framework can be linked to information regarding digital content recorded in the METS file content section. The file content section may be used to record links to content files residing externally to the METS file, or digital content may be Base64 encoded and placed within the file content section itself.[3] The file content section also allows for a certain amount of file-specific technical metadata, such as a checksum, to be recorded for each file, and allows for files to be grouped together into sets. If we scanned page images for a book such as the *Hunting of the Snark*, for example, we might create separate file groups for original, high-resolution master scans, derivative medium-resolution images for web display, and thumbnail images. A single entry for an image file intended for web display might resemble the following:

```
<file ID="PAGE96W" MIMETYPE="image/jpeg" SIZE="93456">
   <FLocat LOCTYPE="URL"
      xlink:href="http://etext.lib.virginia.edu/images/
      modeng/public/CarSnar/CarSna10.jpg" />
```

---

[2] I have attached a complete METS file for a text version of Lewis Carroll's Hunting of the Snark as Appendix I, so that readers can get a more complete picture of what a functioning METS file looks like. I am endebted to the University of Virginia Library for making their electronic text version of this material available.

[3] Well-formed XML content can be placed within a METS file without Base64 encoding, but those creating METS objects containing embedded XML need to take care that such content files are, in fact, well-formed. Those trying to archive web sites using METS, for example, would be well advised to check to see whether pages advertising themselves as XHTML are in fact valid XHTML before attempting to place them within in METS wrapper without using Base64 encoding.

```
</file>
```

with the <file> element possessing an ID attribute which allows this particular file to be referenced from elsewhere in the METS document, and the subsidiary <FLocat> (file location) element recording a URL from which the actual content can be retrieved.

When a content file section for a METS document has been completed, the various divisions within the structural map can then be linked with their corresponding content files. If we assume file elements like the above with ID attribute values of PAGE93W, PAGE94W, PAGE95W and PAGE96W, the <div> element for the eighth fit in Lewis Carroll's work could be modified as follows:

```
<div ORDER="8" TYPE="Fit" LABEL="The Vanishing">
   <fptr>
      <seq>
         <area FILEID="PAGE93W" />
         <area FILEID="PAGE94W" />
         <area FILEID="PAGE95W" />
         <area FILEID="PAGE96W" />
      </seq>
   </fptr>
</div>
```

This encoding uses a file pointer element (<fptr>) to indicate that the content corresponding to this division is a set of files which should be viewed in sequential order (<seq>), with each file identified by an <area> element which indicates the matching <file> element's ID attribute value.[4] Any given <div> may contain multiple

---

[4] METS provides a variety of mechanisms for indicating sequencing which can be confusing at first to those using the standard. The <seq> element indicates a presentation order for a set of files, or portions of files, that correspond to a particular <div> element. The ORDER attribute for the <div> element indicates the logical order of a <div> among its siblings at a particular level of the <div> hierarchy within a <structMap>. This is in turn different from the ORDERLABEL attribute which provides a version of the ORDER value which should be presented to the user. While a bit confusing at first, this variety of sequencing information is necessary for encoding more complex objects.

<fptr> elements to indicate alternative content encodings corresponding with that logical division. One <fptr> could be used to point to master images, while another points to derivatives for web use; if we had a video representation of the story, we could link from the structural map to that as well. Any <fptr> can link to one or more files (or portions of files). If an <fptr> element links to multiple files, it can provide an indication of whether these files should be viewed sequentially or in parallel with one another. METS provides a variety of attributes for the <area> element to allow linking to a portion of a content file (using a time code value to indicate a clip within a video file, for instance, or providing boundary coordinates for an area within a given image). It thus provides a fairly flexible mechanism for recording an abstract intellectual structure for a work and linking that to multiple digital instantiations of the work.

Both the abstract structural elements in the structural map and the content file elements may be linked to a variety of descriptive and administrative metadata. METS is non-prescriptive with respect to descriptive and administrative metadata sets. You are free to use whatever metadata element set you choose; it does not even have to be recorded in XML format, as METS allows you to put either XML or non-XML information within a metadata record.[5] Descriptive and administrative metadata may reside within the METS document, or may be stored externally and referenced via a XLink href attribute. In terms of its handling of descriptive and administrative metadata, METS may be considered an implementation of the Warwick Framework set forth by Lagoze, Lynch and Daniel [12], allowing for multiple and differing metadata packages, both internally and externally referenced, to be brought together in a single framework and linked to the appropriate content.

The final section of METS for conveying information about the digital object's structure is the structural link section (<structLink>). The structural link section employs an XLink extended link in which the "to" and "from" traversal attributes have been constrained to have values of the XML type IDREF. The structural link section is intended to allow for the encoding of links between different <div> elements in the

[5] This facility can be useful for libraries who may simply want to export standard MARC records from their catalog systems and store them in METS documents.

structural map section. This is critical if METS is to be capable of recording non-hierarchical structures which may be present in a digital library object. A common example of where such a facility may be required is web archiving. If you wished to record the hyperlink structures in a website independently of the HTML documents of which the site is composed, the METS structural map is insufficient; while it can record the existence of a page hierarchy for a website, it cannot by itself record the existence of hyperlinks between pages. The structural link section of METS serves as an XLink linkbase, and provides the needed facilities for recording the structure of hyperlinked objects.

The behaviorial metadata section of a METS document records the existence and location of software behaviors which may be needed to view or interact with the digital library object or any of its consitutent parts, including its associated content files. For any given behavior, this section records an optional link to an interface definition for the behavior (such as a WSDL[4] file), and a link to the actual executable mechanism. The behavior may also be linked internally to one or more portions of the METS structural map to indicate which portions of the METS document require the use of this behavior. While the ability to specify software behaviors gives a great deal of control over the dissemination of content and allows institutions to fine tune the presentation of a digital library object, as Dushay [6] notes, the static binding of behaviors to digital objects can be problematic from the point of view of those managing a repository of digital library objects. In particular, it can force repository managers to engage in the rather inefficient process of having to modify every single object in their repository bound to a given behavior whenever the behavior is modified, or when they wish to associate a new behavior with a class of objects. Given these problems, and the traditional concern in the SGML and XML world to insure the separation of content from presentation, a certain degree of caution should be exercised in employing the behavioral metadata section of METS.

The final portion of a METS document is the METS header. This section records a minimal amount of metadata needed for the management of METS documents themselves, including the tracking of creators, editors and other agents who may have worked on a METS document. It also records alternative identifiers which may be assigned to a METS document (the primary identifier for a METS document is recorded in the OBJID attribute on the root METS element), as well as a status indicator for the document.

While the basic components of a METS document are fairly simple, the ability to interrelate the various components means that a METS document for even a relatively straightforward digital library object can become quite complicated. Consider the example of a video resource digitized from a Betacam SP tape to a Motion JPEG 2000 file, which is then used to produce DVD and MPEG4 derivatives. Assume the video has four major sections, and that we have a Dublin Core record for each, as well as a MARC record for the video resource in its entirety. We also wish to record technical metadata about the various digital incarnations, as well as digital provenance information describing how we transferred the video from Betacam SP to Motion JPEG 2000. Figure 1 shows the complex web of relationships that a METS document would need to record, even for a fairly simple resource:

## Figure 1

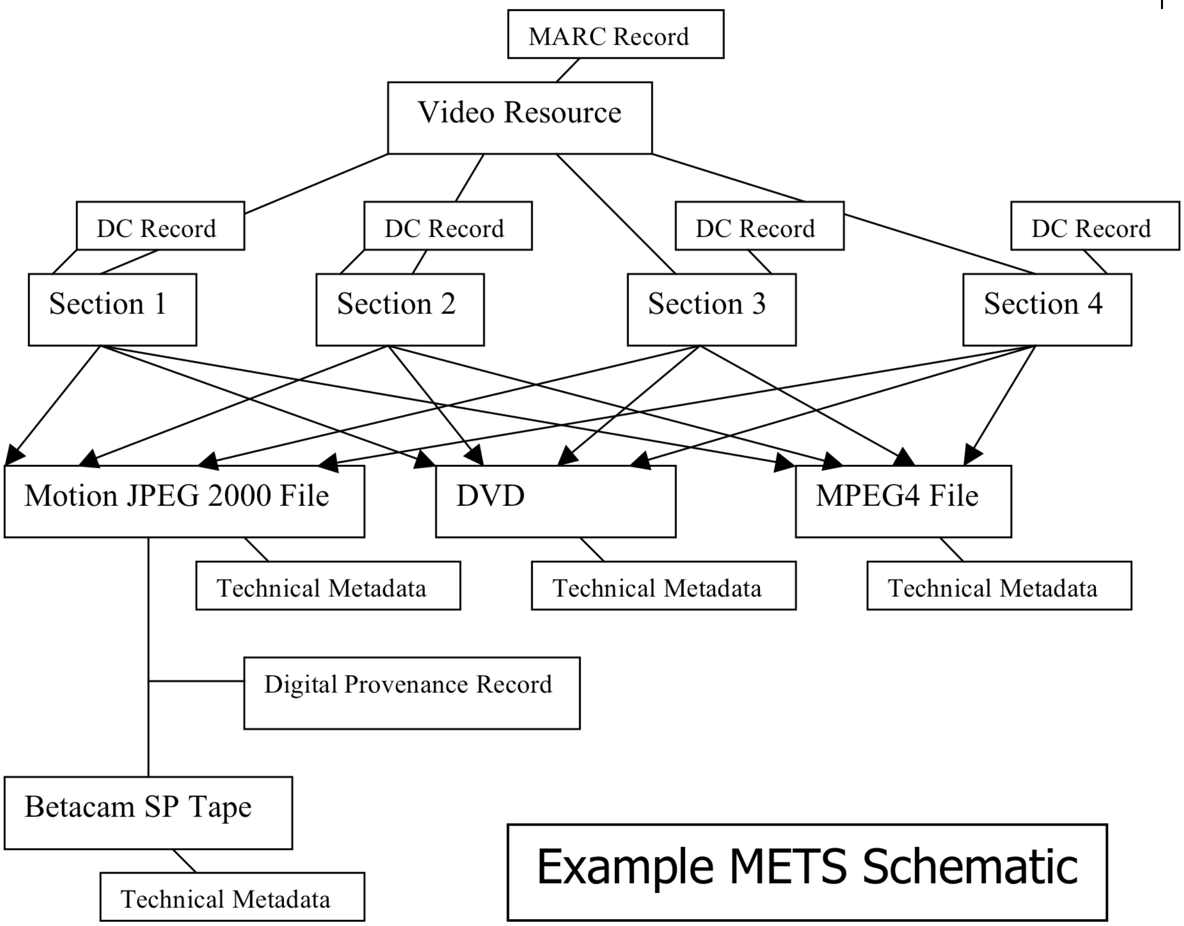


Figure 1 – Relationships tracked by METS for a typical video digital library object

As this case shows, the web of relationships which must be tracked between metadata and content, both digital and analog, can become rather extensive for even the most trivial digital library object. For more complex objects, such as an archived website, the set of relationships which must be tracked and recorded over time can grow quite large. METS, as a document format, is intended to provide a relatively clear point of entry into these complex set of relationships. It is thus very similar in intent to the concept of a 'hub document' put forth in the HyTime standard[9], a document which provides the initial access point for a hyperdocument.

## METS and Interoperability

While METS provides a flexible mechanism for encoding digital library objects, flexibility is often the enemy of interoperability, and METS is no exception to this rule. The potential range of variation in METS documents is extraordinarily high, and the challenges this presents to software developers are considerable. Given METS' flexibility, METS documents created at two different institutions, even for two similar (or identical) objects can be very different. Very basic software operations such indexing, retrieval and display can be difficult to code when the exact nature of the metadata and content to be processed are left as ambiguous as they are in the METS format. A brief review of some of the more problematic areas may help software developers who are working with METS documents, or similarly flexible formats.

The first area of potential difficulty lies in METS' support of the use of arbitrary extension schema. METS does not specify what forms of descriptive or administrative metadata should be used, allowing document encoders to insert whatever metadata schema they wish. Software developers trying to write software to process embedded descriptive metadata within a METS file must be prepared to process MARC/XML, MODS, Dublin Core, EAD, LOM, VRA Core, or any of a variety of other localized descriptive schema. Extension schema for administrative metadata are likely to show even more variability, as there has been less standards development work for administrative metadata to date. Even for a single metadata set, such as Dublin Core, there may be multiple different XML schema that might be used in a METS context. Performing any kind of information processing beyond the most basic keyword extraction is exceptionally difficult when there is no guarantee of what metadata set you'll find, or the exact encoding for a given metadata set.

In addition to allowing the use of arbitrary metadata schema, METS also does not constrain where metadata (or content) may be stored, nor the specific format in which metadata or content is stored. Both metadata and content may be stored internally or externally. Metadata may be in XML or binary form, as may content. There is no guarantee when opening a METS document as to where any of the metadata or content will be located, the format in which any of it will recorded, and in the case of external metadata and content, what the precise access mechanism for retrieving the information will be.

The flexibility that METS provides in terms of metadata schema and the formats employed to record metadata produces another potential difficulty for those trying to process metadata recorded within a METS document; there is no guarantee that any standard rules of description will be applied to metadata recorded within a METS document. While librarians and archivists have developed extensive rules of description, such as the Anglo-American Cataloging Rules [11] and Describing Archives: A Content Standard [14], for use in creating descriptive records, these rules can be difficult to apply outside of the primary metadata set for their communities of practice (MARC and Encoded Archival Description, respectively). Where such rules are applied, content residing within a particular element may be subject to additional processing to extract more detailed information. Name headings in a MARC record formatted according to AACR2, for example, can be relatively easily processed to separate out family names from personal names, even though there are no explicit demarkers to separate them. For other forms of description, however, and for most forms of administrative metadata, there have been no attempts to standardize description. This can make extracting information in a form useful for purposes of indexing and retrieval a very complicated undertaking. The most recent XML schema for Dublin Core, for example, defines the date element as having a data type of string; lacking any rules of description for how a date should be formatted within this element, it is difficult to write software which might extract date information and allow searches such as 'show me works published prior to 1992'.

The METS schema also lacks controlled vocabularies to use in attributes on several elements where they would be extremely helpful to software developers trying to process METS objects. In particular, a controlled vocabulary for use in the TYPE attributes on the root `<mets>` element and on the `<div>` element would assist by allowing software to invoke processing appropriate to the type of object or division, based on a limited set of known types. Unfortunately, there is currently no such vocabulary, and developing a vocabulary comprehensive enough for the range of objects likely to be encoded in METS would be an extremely difficult enterprise.

Perhaps the greatest challenge in writing software to deal with METS documents is the lack of structural constraints METS imposes on a digital object. There are no limitations on the number of descriptive or administrative metadata records. A METS document can also have an unlimited number of structural maps, and there are no restrictions (or even guidelines) on the form a structural map may take, other than those imposed by the METS schema. As a result, a METS document for even so simple and straightforward an object as a book may exhibit radically different structure depending on who is encoding the document and for what purpose.

These problems are not unique to the METS format. To a great extent, they will be shared by any standard for complex digital objects that attempts to include a wide range of potential object formats within its scope. The MPEG-21 Digital Item Declaration Language, for example, has all of the problems set forth above. Descriptors within an MPEG-21 DIDL document (which can be considered analogous to the METS descriptive and administrative metadata sections) may contain text (including XML) or binary information, in any format, or they may reference information living outside of the DIDL document. DIDL imposes no greater degree of structural constraint than the METS format, and is equally impartial with regards to the choice of additional XML schema that may be used to encode information within a DIDL document. The IMS Content Packaging Specification[8] is equally open to the use of any XML schema for encoding metadata within a content package, and places no restrictions on the abstract hierarchical structure that may be defined for a given package. A software developer trying to develop programs to process arbitrary MPEG-21 DIDL, IMS-CP or METS documents will know that they are dealing with an XML document laying out some form of hierarchical structure for the object, and that content files and metadata may be associated with the various parts of that structure, but not a great deal else.

The ability to specify additional constraints on METS documents in order to limit the variation that software developers must confront in writing programs to deal with

METS objects is essential for any practical application of METS.[6] In order to provide a standardized mechanism to express a set of localized constraints for METS documents, the METS editorial board created a separate XML schema for METS profile documents. A METS profile document allows an institution to specify constraints and requirements for METS documents beyond those specified by the METS schema itself. This ability to specify additional constraints on METS documents serves several purposes. It can greatly simplify the job of software developers by narrowing the range of METS documents that their software must support. It enables an institution to achieve a greater degree of consistency in its METS-encoding efforts by allowing it to document local conventions and best practices for document encoding; it also allows an institution to create a specification for locally-acceptable practice with respect to METS document encoding that it may share with other institutions with which it may wish to exchange METS objects, thus promoting interoperability.

A METS profile document allows an institution to document a number of restrictions on, and requirements for, the construction of METS documents. It may restrict the XML schema that may be used in addition to the METS schema for creating METS objects; for example, it may specify that only the Library of Congress MARC/XML schema may be used to record descriptive metadata for METS objects. It may require the application of certain rules of description in various areas of a compliant METS document, such as requiring the use of AACR2 in MARC/XML records. It may require the use of particular controlled vocabularies within a METS document, such as requiring the use of terms from the Getty Thesaurus of Geographic Names for location information. It may impose restrictions on the number of occurences of elements and attributes from the METS schema or other schema within a compliant METS document, such as imposing a limit of a single descriptive metadata record. A profile may establish structural requirements for compliant METS documents. Drawing on our earlier example, it might require that a book of type "Agony" must contain one or more "Fits", or that a movie must contain at

---

[6] Developers working on related document standards have come to the same conclusion. See [1] for information on the application of a Schematron schema as a profiling mechanism for MPEG-21 DIDL documents.

least one scene, and that a scene must contain one or more shots and must not contain subsidiary scenes. It might also impose structural requirements beyond the structural map, such as insisting that document encoders record descriptive metadata within the METS document, and not link to an external metadata record. A profile may also set restrictions on the form of content files, such as insisting that all still image files for a METS object be in the JPEG 2000 format and use the sRGB color space. An institution wishing to exchange METS documents with others needs to be able to provide information regarding the types of METS documents it can accept, and the types of METS documents it can provide. METS profile documents provide a mechanism for documenting and publicizing this information, and thus can assist in establishing the base requirements for interoperability and exchange of digital library materials.

Profiles are no guarantee of interoperability, of course. Institutions can (and doubtless will) adopt METS profiles for local use that will complicate or preclude interoperability with other institutions. By bringing some transparency to local implementation decisions and practice with respect to the METS format, however, METS profiles provide a mechanism for beginning the conversations needed for insuring the ready exchange of complex digital library objects. Given the need for profiling mechanisms across the various complex digital object encoding standards currently in use in research and higher education, further discussion and development of standard profiling mechanisms which could be used in conjunction with a variety of XML encoding formats might be valuable.

## Digital Preservation and Complex Digital Objects

The interoperability problems confronted by those trying to exchange METS documents are significant, but ultimately, these problems can be addressed and solved in a fairly traditional manner: communication and standardization. Librarians have a long history of working together to create the standards they need to exchange information essential to the management and dissemination of materials, and there is no reason to believe they cannot successfully create the standards they need to support the exchange of complex digital objects among themselves and collaborate with other communities of practice to insure the interoperability of content across a variety of systems.

The encoding of complex digital objects presents libraries with other difficulties, however, which are less familiar than the arduous but essential work of creating standards. One of the most difficult of these involves the preservation of complex digital objects. There have been a variety of efforts to develop information models and schema for digital preservation metadata, and many of these efforts are easily translated into an XML format which could be used in conjunction with METS. Most of the digital preservation metadata sets developed in the last several years have a similar framework; they record life-cycle information regarding the curatorial processes to which a particular digital item is subjected. Typically, this information follows an abstract model similar to the following:

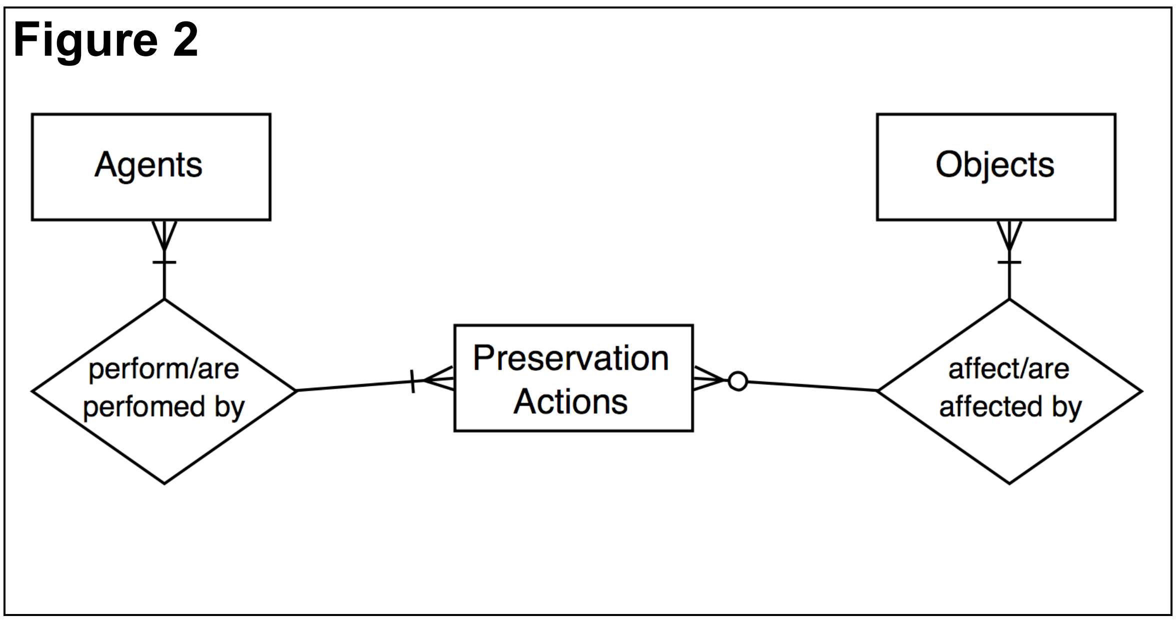


Figure 2: Basic preservation metadata model

Agents managing a repository perform preservation actions on objects, which may in turn result in new objects. These major entities all have a variety of attributes associated with them. Agents will have identifying information such as name, title, contact information, etc. Actions will usually have a date and time, some form of classification to indicate the type of action, and a status marker to indicate the action's outcome; they may also have more specific information regarding a process performed during that action, such as specifying the algorithm used to interpolate missing data in a content object. Objects will have identifying information as well as a variety of technical metadata associated with them. This type of model can capture a variety of life-cycle information for objects; for example, it might indicate that a curator migrated image data stored in a TIFF 6.0 file to a JPEG 2000 file on a particular date using a particular software application. Recording this information documents the custodial history of the data and will assist scholars in the future in determining how accurately the data in front of them reflects the digital information originally captured.

It is the accretive nature of this type of metadata that may ultimately present libraries and archives with a problem in encoding digital library objects. As an example, consider the case experienced at New York University where they are archiving the web sites of organizations active in the labor movement in the United States. As web sites will change over time, for any given organization's web site NYU will typically establish a collection schedule where they will re-harvest the website after a fixed period of time has elapsed. Depending on the degree of change in a site, it might be reharvested anywhere from once a day to once a year.

Small websites have less than a hundred files; the labor sites that NYU have archived typically contain somewhere between 250 and 1000 files. Consider a case of a website such as the Anarchy Archives, with approximately 500 files that we wish to harvest twice a year. We can automate the creation of a METS document to hold these successive harvests, creating a structural map along the following lines:

```
<structMap>
     <div TYPE="Website" LABEL="Anarchy Archives">
          <div TYPE="Harvest" LABEL="June 1, 2003"/>
          <div TYPE="Harvest" LABEL="Dec. 1, 2003"/>
          <div TYPE="Harvest" LABEL="June 1, 2004"/>
          <div TYPE="Harvest" LABEL="Dec. 1, 2004"/>
     </div>
</structMap>
```

Each of the <div> elements for the individual harvests might contain further subdivisions to represent the structure of the website at a particular capture date. The organization of the file content section could take several different forms; the simplest approach, and the easiest to program, would be to simply harvest the entire site each time and create a separate file group for each harvest. A more sophisticated approach might harvest the entire site at the first harvest, and only harvest changed files and additions in later harvests; in that case, a structural map for a later harvest would have links to files from the original harvest as well as to files added to the site or modified since that harvest.

Suppose that the site, when initially captured, contained a number of GIF image files, and that we later decide, based on lessening support for the GIF format, to migrate these files to the JPEG2000 format. Each of the new files will require new technical metadata, although some of it (color space, file format, bit depth) can be stored in a single technical metadata record to which all of the new image files are linked. Each new file will also require a digital preservation record that associates it with the prior version of the file. A decision will need to be made at this point whether the original file will be removed when a new file is created, or whether it will be preserved, but in either case, structural maps that pointed to the old version of the file will need to be updated to point to the new version.

It is possible to accomplish the above using a single METS file, and to some degree it might be easier to manage all this information if it is kept in a single location. But this approach runs a danger of falling prey to a phemonenon which we might call 'the pearl that killed the oyster' syndrome. While it is feasible to keep adding more and more data and metadata to a single METS document, it is not particularly practical. The continual accretion of information within a single METS file will eventually produce a XML file so large that performance for parsing and other computer operations will become unacceptably slow. While advances in hardware and software may eventually ameliorate these performance problems, today they present a serious obstacle to those needing to add extremely large, complex objects into a digital archive.

Appropriate crafting of METS documents can help overcome these problems. In the web archiving example above, there are two forces driving the accretion of information. One is the need to continually reharvest the site to obtain the latest version; the second is the need to document the custodial chain of history for the information from the site by recording digital preservation information when files are migrated. Solving the first of these problems is relatively easy. METS' structural map permits a <div> element to be linked to content residing in another METS document. Instead of encoding the complete structure of each harvested version of the website in a single METS file, we could use the root structure set forth above, and use the METS Pointer element (<mptr>) to indicate that the complete structure and content for a given harvest division can be found in another METS file. We would thus have a "root" METS file containing structural metadata regarding the existence of various harvests, which would then point to other METS documents for detailed information on each harvest.

```
<structMap>
     <div TYPE="Website" LABEL="Anarchy Archives">
          <div TYPE="Harvest" LABEL="June 1, 2003">
               <mptr LOCTYPE="URL"
               xlink:href="http://dlib/metsfile2.xml"
               />
          </div>
          .
          .
     </div>
</structMap>
```

By separating the harvests into external METS files, we can avoid the problem of having a single file for an archived web site accreting a vast amount of information.

This also reduces the second problem somewhat, as the digital preservation information for migrated files will be distributed across several METS files, rather than residing in a single file. However, in the long term, a great deal of digital preservation information might accrete in the METS file for a single harvest. A possible solution to this would be to take the use of the <mptr> element to another level, and use a separate METS document for each file harvested from the web site. When a content file is migrated to a new format, the METS document for the newer version of the file would contain a digital preservation record describing the migration process and providing a link to the METS file for the older version. The structural map for the harvested site would be updated to link to the new versions of the files; those interested in the previous incarnations of the data could follow the chain of digital preservation references back to the original information.

If the content files in this scenario were Base64 encoded and wrapped within their individual METS files, the METS files would become somewhat analogous to the Universal Preservation Format proposed by Shepard and MacCarn[13], a wrapper containing digital 'essence,' technical metadata necessary for decoding that essence, and other descriptive, rights and preservation metadata needed to manage the file. This approach would certainly go some way towards solving the performance problems presented by slowly accreting METS documents. It would also, however, incur some performance penalties of its own. Extracting the file content from the XML wrapper and reversing the Base64 encoding whenever the content of a file needed to be displayed is obviously a slower operation than simply reading the unencoded data file on its own. But for large, structurally complex items, the performance degradation resulting from having to extract and decode the information for individual files might well be balanced out by reducing the amount of XML processing which must occur to render portions of a METS document for display. Some practical research into the performance tradeoffs involved in encoding a complex digital object using a large, monolithic XML file versus decomposing the object into a set of small, discrete XML files would be valuable for those wishing to fine tune their local encoding practices to enhance performance of digital library systems.

## Conclusions

METS provides an open, standardized XML encoding format for storing the descriptive, administrative, structural and behavioral metadata needed to manage complex digital objects. It also provides mechanisms for recording the complex inter-relations of metadata and content for those objects. As the OAIS Reference Model makes clear, managing complex digital objects and insuring their long-term preservation requires a packaging mechanism which can support a large and complex mix of descriptive, technical, rights, digital preservation and structural metadata. METS was designed to provide digital libraries with a practical and flexible packaging mechanism for digital objects to support their long-term preservation and promote the interoperability of digital library objects between different repository systems.

While METS has gone some distance towards achieving these design goals, it is not in itself a guarantee of interoperability, and there are some obvious practical difficulties in using METS for the long-term preservation of digital objects. These issues are not unique to METS, and indeed, many of those working on similar standards for complex digital content for use in research and higher education are encountering the same issues that digital library developers employing METS are experiencing. While it is unfortunate that we do not appear to have solved all of the world's problems with respect to complex object encoding, there is a unique opportunity at the moment to collaborate with others outside the library community to try to further refine our standards for encoding complex digital objects, and to share the knowledge we have gained in the development and application to METS with others who are confronting the same issues.

## Appendix I

```
<?xml version="1.0" encoding="UTF-8"?>
<mets xmlns="http://www.loc.gov/METS/"
     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
     xmlns:xlink="http://www.w3.org/1999/xlink"
     xmlns:dc="http://purl.org/dc/elements/1.1/"
     xmlns:UVALocal="http://www.virginia.edu/local"
     xsi:schemaLocation="http://www.loc.gov/METS/
file:/Users/jeromemcdonough/Documents/METS/mets.v1-5.xsd">
    <dmdSec ID="DM1">
        <mdWrap MDTYPE="DC">
            <xmlData>
                <dc:dc>
                    <dc:title>The Hunting of the Snark: an Agony
in Eight Fits</dc:title>
                    <dc:creator>Carroll, Lewis</dc:creator>
                    <dc:publisher>Electronic Text Center,
University of Virginia</dc:publisher>
                    <dc:date>1876</dc:date>
                    <dc:type>text</dc:type>
                    <dc:language>en</dc:language>
                    <dc:rights>copyright 1998 University of
Virginia Library</dc:rights>
                </dc:dc>
            </xmlData>
        </mdWrap>
    </dmdSec>
```

<amdSec>
<rightsMD ID="R1">
<mdWrap MDTYPE="OTHER" OTHERMDTYPE="local rights statement">
<xmlData>
<UVALocal:rights>

```
basis and we are actively working our way through the public
ASCII texts on the Internet whose provenance can be determined.
                    </UVALocal:rights>
                </xmlData>
            </mdWrap>
        </rightsMD>
        <digiprovMD ID="DP1">
            <mdWrap MDTYPE="OTHER" OTHERMDTYPE="local digital
provenance statement">
                <xmlData>
                    <UVALocal:provenance>
Creation of machine-readable version, depositing the text on
Project Gutenberg.

Conversion to TEI-conformant markup: University of Virginia
Library Electronic Text Center ca. 45 kilobytes -- rounded up to
the neareste 5KB.

This version available from the University of Virginia Library,
Charlottesville, Va.
                    </UVALocal:provenance>
                </xmlData>
            </mdWrap>
        </digiprovMD>
    </amdSec>
    <fileSec>
        <fileGrp>
            <file ID="F1" ADMID="R1 DP1"><FLocat LOCTYPE="URL"
xlink:href="http://etext.lib.virginia.edu/etcbin/toccer-
new2?id=CarSnar.sgm&amp;images=images/modeng&amp;data=/texts/engl
ish/modeng/parsed&amp;tag=public&amp;part=1&amp;division=div1"/><
/file>
            <file ID="F2" ADMID="R1 DP1"><FLocat LOCTYPE="URL"
xlink:href="http://etext.lib.virginia.edu/etcbin/toccer-
new2?id=CarSnar.sgm&amp;images=images/modeng&amp;data=/texts/engl
ish/modeng/parsed&amp;tag=public&amp;part=2&amp;division=div1"/><
/file>
            <file ID="F3" ADMID="R1 DP1"><FLocat LOCTYPE="URL"
xlink:href="http://etext.lib.virginia.edu/etcbin/toccer-
new2?id=CarSnar.sgm&amp;images=images/modeng&amp;data=/texts/engl
ish/modeng/parsed&amp;tag=public&amp;part=3&amp;division=div1"/><
/file>
            <file ID="F4" ADMID="R1 DP1"><FLocat LOCTYPE="URL"
xlink:href="http://etext.lib.virginia.edu/etcbin/toccer-
new2?id=CarSnar.sgm&amp;images=images/modeng&amp;data=/texts/engl
ish/modeng/parsed&amp;tag=public&amp;part=4&amp;division=div1"/><
/file>
            <file ID="F5" ADMID="R1 DP1"><FLocat LOCTYPE="URL"
xlink:href="http://etext.lib.virginia.edu/etcbin/toccer-
new2?id=CarSnar.sgm&amp;images=images/modeng&amp;data=/texts/engl
ish/modeng/parsed&amp;tag=public&amp;part=5&amp;division=div1"/><
/file>
```

```
            <file ID="F6" ADMID="R1 DP1"><FLocat LOCTYPE="URL"
xlink:href="http://etext.lib.virginia.edu/etcbin/toccer-
new2?id=CarSnar.sgm&amp;images=images/modeng&amp;data=/texts/engl
ish/modeng/parsed&amp;tag=public&amp;part=6&amp;division=div1"/><
/file>
            <file ID="F7" ADMID="R1 DP1"><FLocat LOCTYPE="URL"
xlink:href="http://etext.lib.virginia.edu/etcbin/toccer-
new2?id=CarSnar.sgm&amp;images=images/modeng&amp;data=/texts/engl
ish/modeng/parsed&amp;tag=public&amp;part=7&amp;division=div1"/><
/file>
            <file ID="F8" ADMID="R1 DP1"><FLocat LOCTYPE="URL"
xlink:href="http://etext.lib.virginia.edu/etcbin/toccer-
new2?id=CarSnar.sgm&amp;images=images/modeng&amp;data=/texts/engl
ish/modeng/parsed&amp;tag=public&amp;part=8&amp;division=div1"/><
/file>
        </fileGrp>
    </fileSec>
    <structMap>
        <div DMDID="DM1" ORDER="1" TYPE="Agony" LABEL="Hunting of
the Snark">
            <div ORDER="1" TYPE="Fit" LABEL="The Landing">
                <fptr FILEID="F1"></fptr>
            </div>
            <div ORDER="2" TYPE="Fit" LABEL="The Bellman's
Speech">
                <fptr FILEID="F2"></fptr>
            </div>
            <div ORDER="3" TYPE="Fit" LABEL="The Baker's Tale">
                <fptr FILEID="F3"></fptr>
            </div>
            <div ORDER="4" TYPE="Fit" LABEL="The Hunting">
                <fptr FILEID="F4"></fptr>
            </div>
            <div ORDER="5" TYPE="Fit" LABEL="The Beaver's
Lesson">
                <fptr FILEID="F5"></fptr>
            </div>
            <div ORDER="6" TYPE="Fit" LABEL="The Barrister's
Dream">
                <fptr FILEID="F6"></fptr>
            </div>
            <div ORDER="7" TYPE="Fit" LABEL="The Banker's Fate">
                <fptr FILEID="F7"></fptr>
            </div>
            <div ORDER="8" TYPE="Fit" LABEL="The Vanishing">
                <fptr FILEID="F8"></fptr>
            </div>
        </div>
    </structMap>
</mets>
```

# References

[1] Bekaert J, Hochstenbach P, Van de Sompel H (2003) Using MPEG-21 DIDL to Represent Complex Digital Objects in the Los Alamos National Laboratory Digital Library. D-Lib Magazine 9(11). Retrieved on Nov. 22, 2004 from http://www.dlib.org/dlib/november03/bekaert/11bekaert.html

[2] Biron PV, Malhotra A (2004) XML Schema Part 2: Datatypes. 2nd Edition. Cambridge, MA: World Wide Web Consortium. Retrieved on Nov. 22, 2004 from http://www.w3.org/TR/xmlschema-2/

[3] Carroll L (1891) The Hunting of the Snark: An Agony in Eight Fits. New York: MacMillan and Co. Retrieved on Nov. 22, 2004 from http://etext.lib.virginia.edu/etcbin/toccer-new2?id=CarSnar.sgm&images=images/modeng&data=/texts/english/modeng/parsed&tag=public&part=all

[4] Christensen E, Curbera F, Meredith G, Weerawarana S (2001) Web Services Description Language (WSDL) 1.1: W3C Note 15 March 2001. Cambridge, MA: World Wide Web Consortium. Retrieved on Nov. 22, 2004 from http://www.w3.org/TR/wsdl

[5] Consultative Committee for Space Data Systems (2003) Space Data and Information Transfer Systems -- Open Archival Information System -- Reference Model (ISO 14721:2003). Geneva: International Organization for Standardization.

[6] Dushay N (2002) Localizing Experience of Digital Content via Structural Metadata. Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries, Portland, Oregon, USA, July 13-17, 2002. Downloaded Nov. 22, 2004 from http://portal.acm.org/citation.cfm?id=544220.544274

[7] Hurley B, Price-Wilkin J, Proffitt M, Besser H (1999) The Making of America II Testbed Project: A Digital Library Service Model. Washington, DC: The Digital Library Federation / Council on Library and Information Resources. Retrieved on Nov. 22, 2004 from http://www.clir.org/pubs/reports/pub87/pub87.pdf

[8] IMS Global Learning Consortium (2004) IMS Content Packaging – Version 1.1.4 – Final Specification. Burlington, MA: Author.

[9] International Organization for Standardization (1997) Information processing -- Hypermedia/Time-based Structuring Language (HyTime) - 2d edition. ISO/IEC 10744:1997. Goldfarb CF, Newcomb SR, Kimber WE, Newcomb PJ (eds). Geneva: Author.

[10] International Organization for Standardization (2003). Information technology – Multimedia framework (MPEG-21) – Part 2: Digital item Declaration. ISO /IEC 21000-2. Geneva: Author.

[11] Joint Steering Committee for Revision of AACR, a committee of the American Library Association, the Australian Committee on Cataloguing, the British Library, the Canadian Committee on Cataloguing, the Library Association, the Library of Congress (1988) Anglo-American Cataloguing Rules – 2nd ed – 1988 revision. Gorman M, Winkler PW (eds). Ottawa : Canadian Library Association ; London : Library Association Publishing ; Chicago : American Library Association.

[12] Lagoze C, Lynch CA, Daniel, Jr., R. (1996) The Warwick Framework: A Container Architecture for Aggregating Sets of Metadata. Ithaca, NY: Cornell University Library. Downloaded on Nov. 22, 2004 from http://techreports.library.cornell.edu:8081/Dienst/UI/1.0/Display/cul.cs/TR96-1593?abstract=

[13] Shepard T, MacCarn D (2001) The Universal Preservation Format: A Recommended Practice for Archiving Media and Electronic Records. Boston, MA: WGBH Educational Foundation. Retrieved on Nov. 22, 2004 from http://info.wgbh.org/upf/pdfs/20010818-UPF_RP.pdf

[14] Society of American Archivists (2004) Describing Archives: A Content Standard. Chicago: Author.

[15] The TEI Consortium (2004) TEI P4: Guidelines for Electronic Text Encoding and Interchange. Sperberg-McQueen CM, Burnard L (eds). Bergen, Norway: Author.

[16] Thompson HS, Beech D, Maloney M, Mendelsohn N (2004) XML Schema Part 1: Structures. 2nd Edition. Cambridge, MA: World Wide Web Consortium. Retrieved on Nov. 22, 2004 from http://www.w3.org/TR/xmlschema-1/