Metadata Encoding & Transmission Standard

Home page:

http://www.loc.gov/standards/mets/

What is Metadata?

- It is “data about data”
- The focus of "aboutness" is to enhance the use of the source data ...
- Metadata from this perspective is any information which supports the effective use of data, including information which can facilitate data management (e.g., data authentication, data sequence, data type, key field indicators), data access (e.g., range, report parameter) and data analysis (e.g., format for data mining, visualization)

Quoted in http://www.scils.rutgers.edu/~sypark/asis.html

Not again?!

- Yet another metadata encoding scheme!
- Extant schemas include:
  - MARC (original LOC scheme dating from early 80's)
  - Dublin Core (original flavor and new qualified version)
  - BibTeX/TeX/LaTeX
  - XML and its children (RDF, XHTML and WSDL, for example)
  - NISOIMG, VRA Core, (NISO Technical Metadata for Digital Still Images), LC-AV (Library of Congress Audiovisual Metadata), etc. etc. etc.
- Eventually meta-meta-metadata encoding → *meta*^{n}?
- So why bother?

Role of Metadata: Traditional Library

- Bibliographic Role (library catalogue)
  - Search/locate an item
  - List/organize collection of items
  - Aid in item selection from among numerous options

Role of Metadata: Digital Library

- Card catalogue replaced with more flexible electronic system
- Information objects can be located in more than one place, but still serve to enumerate the collection
- Digital format facilitates item selection process

A richer, more flexible and adaptable use of the original card catalog concept, but fundamentally the same

*The **key** to digital library access **is** the **metadata***

Goal of METS

To provide a flexible mechanism:

- for encoding descriptive, administrative, and structural metadata for a digital library object
- for expressing the complex links between these various forms of metadata
- which offers a useful standard for the exchange of digital library objects between repositories
- enable the association of a digital object with behaviors or services

7 Basic METS Components

1. METS Header
2. Descriptive Metadata
3. Administrative Metadata
4. File Section
5. Structural Map
6. Structural Links
7. Behavior

METS Header Example

<metsHdr CREATEDATE="2003-07-04T15:00:00" RECORDSTATUS="Complete"> <agent ROLE="CREATOR" TYPE="INDIVIDUAL"> <name>Jerome McDonough</name> </agent> <agent ROLE="ARCHIVIST" TYPE="INDIVIDUAL"> <name>Ann Butler</name> </agent> </metsHdr>

Descriptive Metadata Examples

<dmdSec ID="dmd001"> <mdRef LOCTYPE="URN" MIMETYPE="application/xml" MDTYPE="EAD" LABEL="Berol Collection Finding Aid">urn:x-nyu:fales1735</mdRef> </dmdSec>

<dmdSec ID="dmd002"> <mdWrap MIMETYPE="text/xml" MDTYPE="DC" LABEL="Dublin Core Metadata"> <xmlData> <dc:title>Alice's Adventures in Wonderland</dc:title> <dc:creator>Lewis Carroll</dc:creator> <dc:date>between 1872 and 1890</dc:date> <dc:publisher>McCloughlin Brothers</dc:publisher> <dc:type>text</dc:type> </xmlData> </mdWrap> </dmdSec>

<dmdSec ID="dmd003"> <mdWrap MIMETYPE="application/marc" MDTYPE="MARC" LABEL="OPAC Record"> <binData>MDI0ODdjam0gIDIyMDA1ODkgYSA0NU0wMDAxMDA...(etc.) </binData> </mdWrap> </dmdSec>

Administrative Metadata Example

```
<amdSec ID="AMD001">
    <mdWrap MIMETYPE="text/xml" MDTYPE="NISOIMG" LABEL="NISO
  Img. Data">
      <niso:MIMEtype>image/tiff</niso:MIMEtype>
      <niso:Compression>LZW</niso:Compression>
      <niso:PhotometricInterpretation>8</niso:PhotometricInterpretation>
      <niso:Orientation>1</niso:Orientation>
      <niso:ScanningAgency>NYU Press</niso:ScanningAgency>
    </mdWrap>
  </amdSec>
```

A <file> element within a <fileGrp> might then identify this administrative metadata as pertaining to the file it identifies by using an ADMID attribute to point to this <amdSec> element:

```
<file ID="FILE001" ADMID="AMD001">
    <FLocat LOCTYPE="URL">http://dlib.nyu.edu/press/testimg.tif</FLocat>
  </file>
```

File Section Example

<fileSec> <fileGrp ID="VERS1"> <file ID="FILE001" MIMETYPE="application/xml" SIZE="257537" CREATED="2001-06-10">
<FLocat LOCTYPE="URL">http://dlib.nyu.edu/tamwag/beame.xml</FLocat>
</file> </fileGrp>

<fileGrp ID="VERS2"> <file ID="FILE002" MIMETYPE="audio/wav" SIZE="64232836" CREATED="2001-05-17" GROUPID="AUDIO1">
<FLocat LOCTYPE="URL">http://dlib.nyu.edu/tamwag/beame.wav</FLocat>
</file> </fileGrp>

<fileGrp ID="VERS3" VERSDATE="2001-05-18"> <file ID="FILE003" MIMETYPE="audio/mpeg" SIZE="8238866" CREATED="2001-05-18" GROUPID="AUDIO1">
<FLocat LOCTYPE="URL">http://dlib.nyu.edu/tamwag/beame.mp3</FLocat>
</file> </fileGrp> </fileSec>

Structural Map Example

<structMap TYPE="logical">
 <fptr FILEID="FILE001"> <area FILEID="FILE001" BEGIN="INTVWBG" END="INTVWND" BETYPE="IDREF" /> </fptr> <fptr FILEID="FILE002"> <area FILEID="FILE002" BEGIN="00:00:00" END="00:01:47" BETYPE="TIME" /> </fptr> <fptr FILEID="FILE003"> <area FILEID="FILE003" BEGIN="00:00:00" END="00:01:47" BETYPE="TIME" /> </fptr>

 <fptr FILEID="FILE001"> <area FILEID="FILE001" BEGIN="FHBG" END="FHND" BETYPE="IDREF" /> </fptr> <fptr FILEID="FILE002"> <area FILEID="FILE002" BEGIN="00:01:48"END="00:06:17" BETYPE="TIME" /> </fptr> <fptr FILEID="FILE003"> <area FILEID="FILE003" BEGIN="00:01:48" END="00:06:17" BETYPE="TIME" /> </fptr>

 <fptr FILEID="FILE001"> <area FILEID="FILE001" BEGIN="TUBG" END="TUND" BETYPE="IDREF" /> </fptr> <fptr FILEID="FILE002"> <area FILEID="FILE002" BEGIN="00:06:18" END="00:10:03" BETYPE="TIME" /> </fptr> <fptr FILEID="FILE003"> <area FILEID="FILE003" BEGIN="00:06:18" END="00:10:03" BETYPE="TIME" /> </fptr>
 </structMap>

Structural Links Example

 <fptr FILEID="HTMLF1"/> <fptr FILEID="JPGF1"/>   <fptr FILEID="HTMLF2"/>

If you wished to indicate that the image file in the contained with the first page is hyperlinked to the HTML file in the second page , you would have a <smLink> element within the <structLink> section of the METS document as follows:

<smLink from="IMG1" to="P2" xlink:title="Hyperlink from JPEG Image on Page 1 to Page 2" xlink:show="new" xlink:actuate="onRequest" />

Behavior Section Example

Digital object behaviors can be implemented as linkages to distributed web services as in the following example from the Mellon Fedora project.

<METS:behavior ID="DISS1.1" STRUCTID="S1.1" BTYPE="uva-bdef:stdImage" CREATED="2002-05-25T08:32:00" LABEL="UVA Std Image Disseminator" GROUPID="DISS1" ADMID="AUDREC1"> <METS:interfaceDef LABEL="UVA Standard Image Behavior Definition" LOCTYPE="URN" xlink:href="uva-bdef:stdImage"/> <METS:mechanism LABEL="A NEW AND IMPROVED Image Mechanism" LOCTYPE="URN" xlink:href="uva-bmech:BETTER-imageMech"/> </METS:behavior>

Conclusion & Opinions

- There are numerous metadata schemes out there designed to solve the digital library cataloging/access/preservation problem
- METS is a LOC attempt to enable these to work together in an 'envelope' of uber-meta-data
  - It can stand alone as a metadata scheme –or-
  - It can wrap existing schemes within (including MARC)
  - Acknowledges possible inclusion of the digital object within the metadata records (others often do not)
  - Addresses common DL issues not handled by other many other schemes (behaviors, IPR, services)
- The good news: It is complex, full-featured, extensible
- The bad news: It is complex, full-featured, extensible
- *Opinion:* DL cataloging is still looking for its silver bullet, equivalent to the traditional library card catalog but with the added richness information retrieval that digital media can have.