

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

8-6-2014

METS For The Cultural Heritage Community: A Literature Review

Sharon Cheslow

San Jose State University, sharoncheslow@gmail.com

Follow this and additional works at: <http://digitalcommons.unl.edu/libphilprac>



Part of the [Digital Humanities Commons](#), and the [Library and Information Science Commons](#)

Cheslow, Sharon, "METS For The Cultural Heritage Community: A Literature Review" (2014). *Library Philosophy and Practice (e-journal)*. 1162.

<http://digitalcommons.unl.edu/libphilprac/1162>

METS for the Cultural Heritage Community: A Literature Review

Sharon Cheslow

San José State University

August 6, 2014

Abstract

METS (Metadata Encoding and Transmission Standard) is an XML-based, data communication standard used for digital collections in cultural heritage institutions, including libraries, archives, and museums, and maintained by the Library of Congress. Recent articles have been written for those in the cultural heritage community who may find METS beneficial. Even so, the uses of METS are still being discovered in terms of best practices and interoperability. One of the main issues with METS is how it can be used with external schemas such as MODS, PREMIS, or Dublin Core. This paper includes a brief description of METS as a wrapper with external metadata schemas, followed by a literature review focusing on METS' development since 2001, and its recent uses with external schemas.

Keywords

METS, MODS, PREMIS, XML, metadata, Library of Congress, literature review, digital libraries, digital objects, interoperability

Introduction

METS, which stands for Metadata Encoding and Transmission Standard, is a data communication standard used for digital collections in cultural heritage institutions, including libraries, archives, and museums (Cantara, 2005; Cundiff, 2004; Digital Library Foundation, 2010; Elings & Waibel, 2007; Guenther & McCallum, 2003, Mazé, 2012; NISO, 2004). The METS schema is an open standard used for both storing and transmitting data for digital resources. It was developed by the Digital Library Foundation a little over ten years ago, and is currently maintained by the Library of Congress (Library of Congress, 2012, METS Web Site). METS is often used to encode metadata for digital audio, digital video, or digital images. This paper will first provide a brief description of METS as a wrapper with external metadata schemas. The literature review section will then focus on METS' development since 2001, its recent uses with external schemas such as MODS and PREMIS, and its use with digital audio and visual metadata records.

METS and External Schemas

A METS record is comprised of a wrapper. The wrapper begins and ends with METS tags, and surrounds the external schema with which it is used. The METS wrapper has a basic structure, with up to seven major subsections: a METS Header (metsHDR), a Descriptive Metadata Section (dmdSec), an Administrative Metadata Section (amdSec), a File Section (fileSec), a Structural Map (structMap), Structural Links (structLink), and a Behavior Section (behaviorSec). The METS document must have at minimum the Structural Map (Cundiff, 2004; Digital Library Foundation, 2010). Each of these subsections has elements that provide the means for describing in detail the digital objects. The Structure Map or Structural Map defines a

hierarchical structure so users of the digital library object can navigate through it via METS pointers (LC, 2011, METS: An Overview and Tutorial).

“External schemas are schemas that define an XML vocabulary and syntax appropriate for use in conjunction with METS in its descriptive and administrative metadata contexts” (Library of Congress, 2011, METS Extenders). METS is displayed using XML (Extensible Markup Language), and it is especially useful as a wrapper with MODS (Metadata Object Description Schema), Dublin Core, and PREMIS (Preservation Metadata: Implementation Strategies). MODS and Dublin Core are descriptive metadata schemas used for identifying digital resources, and PREMIS is an administrative metadata schema used for the preservation of digital resources (Bolin, February 2013).

METS Literature Review

Library of Congress METS Web Site

The intended audience for METS is generally the cultural heritage community, including digital collections in libraries, archives, and museums. Since the METS standard is maintained in the Network Development and MARC Standards Office of the Library of Congress, the Library of Congress METS web site is an invaluable resource for digital libraries and collections with digital objects. The METS web site offers an abundance of resources, which are well organized and easy to navigate, including a basic introductory article (Library of Congress, 2011, METS: An Overview and Tutorial).

According to the Library of Congress (2012, METS Web Site), “The METS schema is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema language of the World Wide Web Consortium.” The Library of Congress offers information on external schemas that can be used

with METS. The following descriptive metadata schemas have been endorsed by the METS Editorial Board for use with METS: Dublin Core, MODS, MARCXML, and VRA Core. METS has also been endorsed for use with the following administrative metadata schemas: textMD (Schema for Technical Metadata for Text), and MIX (NISO Technical Metadata for Digital Still Images). There are also specific audio and video metadata schemas that are compatible with METS, including audioMD and videoMD. Finally, METS is endorsed for use with PREMIS for preservation metadata.

The Library of Congress provides tools on integrating these schemas, and these tools support external metadata creation for use with METS (Library of Congress, 2012, METS Tools and Utilities). The Library of Congress' METS web site also includes links to examples of METS documents for the following: bibliographic records, page turners, maps, various types of images, PDFs, sheet music and scores, sound recordings, realia, serials, and video (Library of Congress, 2011, Example Documents). Overall, the Library of Congress METS web site is an important resource for those seeking information on METS.

Articles/Reports

What does scholarly and professional literature say about METS' development since 2001 for the cultural heritage community, including its uses with audio and visual resources? There is not an abundance of literature available on METS, since it is a relatively recent schema. Most of the material comes from the Library of Congress, since they maintain the metadata standard. There are, however, some excellent introductory scholarly articles from library and archival journals that provide an analysis of how METS can be used with digital objects. In addition, there are more technically advanced recent articles dealing with digital libraries and

digital collections in cultural heritage institutions that discuss METS. This section will discuss relevant articles in chronological order.

In 2003 Rebecca Guenther and Sally McCallum wrote an article for the *Bulletin of the American Society for Information Science and Technology*, when they were employed with the Network Development and MARC Standards Office at the Library of Congress. Their article “New Metadata Standards for Digital Resources: MODS and METS” is a great introduction to the early uses of these metadata schema for use in digital libraries. The authors explain the Library of Congress developed MODS, an XML schema that includes a subset of MARC 21 data elements, in 2002. Guenther and McCallum (2003) state METS “grew out of several experimental 1990s digital projects, [and in] 2001, the Digital Library Federation convened a meeting of experts from several projects to evaluate what had been learned...Out of that meeting came the idea for METS, an XML document that packages the metadata associated with a digital resource...In a little over a year the METS XML schema was developed, a maintenance structure set up and experimentation worldwide began” (METS section). This article appears to be one of the first written about METS, two years after the idea was developed and when experimentation using MODS with METS was just beginning. One such early experimental project the authors mention is the Library of Congress' AudioVisual Prototyping Project. At the time this article was written, Guenther and McCallum pointed out the Library of Congress already had over 7 million digital objects. They believed METS was essential for a digital repository. Since the article was co-written by staff at the Library of Congress, it is very useful. In addition, it is still relevant.

What exactly is METS in terms of metadata schemas? Ten years ago NISO published *Understanding Metadata*, an invaluable resource available as an online PDF (National Information Standards Organization, 2004). The publication helps readers grasp what metadata

is, through descriptions of different types of metadata. It is a technical document, but not too difficult to read. There is a two-page section on METS that discusses its history, characteristics, and elements. The METS section also briefly describes an example of the use of METS with a digitized oral history collection, although it doesn't provide an example record.

Morgan Cundiff's 2004 article, "An Introduction to the Metadata Encoding and Transmission Standard (METS)" in *Library Hi Tech*, provides a helpful overview of METS for library and technical staff encountering this schema for the first time. As of 2004, Cundiff was Senior Standards Specialist in the Network Development and MARC Standards Office. Her article contains a brief history of METS' development, a section covering the basic structure and content of METS documents, and a discussion of several issues relevant to its implementation and continuing development, including external schema. Cundiff (2004) points out that the 2001 Digital Library Federation meeting, which was directed by Jerome McDonough of New York University and described in Guenther and McCallum's article, led to a consensus in the library community that METS should be an XML schema. It was decided that descriptive and administrative metadata could be expressed using vocabularies from external schema. Cundiff also points out METS provides a method for aggregating all the metadata that can be used with a digital object. Although the article is very technical and detailed, it is useful in that it gives specific examples of how METS is used with XML in metadata records.

"METS: The Metadata Encoding and Transmission Standard" is a 2005 article by Lynn Cantara, published in a special metadata issue of *Cataloging & Classification Quarterly*, Cantara was a metadata librarian at Kelvin Smith Library at Case Western Reserve University. Like Cundiff's article, this one is very technical and detailed, with good examples. The author describes in depth the different subsections of the METS wrapper (Cantara, 2005). The article

doesn't give as much introductory information as Cundiff's, and would be a better article for someone already familiar with METS. This would be a good follow-up to Cundiff's article or the Library of Congress' "METS Overview" article.

According to Mary Eilings and Günter Waibel (2007), METS is similar to other XML-based data standards used in the cultural heritage community, such as EAD, since XML encodes the data fields and their values. Their article "Metadata For All: Descriptive Standards and Metadata Sharing Across Libraries, Archives, and Museums" demystifies metadata by examining different standards. They argue that applying standards by material type, and not by type of institution, could lead to greater interoperability within the cultural heritage community.

The American Archivist has many articles on metadata standards for digital objects, but none specifically on METS. However, Jenn Riley and Keley Shepherd's 2009 article "A Brave New World: Archivists and Shareable Descriptive Metadata" is an excellent exploration into the current emphasis on sharability for metadata in the cultural heritage community, with a brief mention of METS. The article mostly goes into depth in discussing EAD and DACS, and also looks at their relation to MODS. The authors point out that Archivists' Toolkit supports export of a METS wrapper with either MODS or Dublin Core. The authors point out that although metadata standards differ among different communities (libraries, archives, and museums), sharing standards can be beneficial in helping to promote common goals (Riley & Shepherd, 2009).

Many of the articles on METS between 2008-2010 were published in *D-Lib Magazine*, an online magazine published by D-Lib Alliance that "provides much of the core literature on digital libraries" (Corporation for National Research Initiatives, 2013). The following paragraph discusses four articles published in *D-Lib Magazine*.

Three of these four *D-Lib Magazine* articles are on the use of METS as a wrapper or container for PREMIS implementation; the fourth looks at METS with both PREMIS and MODS. “The Australian METS Profile – A Journey about Metadata” is a very technical article that chronicles research on packaging and exchanging digital content, based on an Australian METS Profile that the National Library of Australia registered with the Library of Congress (Pearce, Pearson, Williams, & Yeadon, 2008). The authors, all with the National Library of Australia, chose METS extended by PREMIS at the outset in order to discover a generic METS profile that could be shared. Rebecca Guenther’s 2008 article “Buzzwords: Flexibility vs. Interoperability When Implementing PREMIS in METS” can be seen as a more in-depth follow-up to her 2003 article on MODS and METS mentioned previously. In this more recent article, Guenther explores both PREMIS and METS, and discusses guidelines on their use together (Guenther, 2008). Well-written and informative, the article offers insight and technical details into how PREMIS can be used in METS administrative metadata subsections. Especially helpful is a section on redundancies between PREMIS and METS. “Using METS, PREMIS and MODS for Archiving eJournals” offers a report on the use of these standards for structural, preservation, and descriptive metadata, in order to guide institutions that are developing archival digital repositories for e-journals (Dappert, & Enders, 2008). Finally, “A Checklist and a Case for Documenting PREMIS-METS Decisions in a METS Profile” introduces a checklist of key PREMIS-METS issues, each illustrated with examples from Library of Congress METS profiles (Vermaaten, 2010). The author, Sally Vermaaten, proposed this checklist in order to increase efficient documentation and to contribute to best practices for shared metadata, in order to foster preservation and interoperability. According to Vermaaten (2010), “Shared PREMIS-METS practices would also help foster the development of tools for metadata manipulation and creation

that could be used across projects and repositories” (Benefits of Shared Practices section). These articles all offer useful information, guidelines, and best practices for using METS with external schemas for digital objects in cultural heritage institutions.

An in-depth *Library Quarterly* article from 2010 on using METS with external schemas for digital image objects in libraries, archives, and museums is “The Convergence of Information Technology, Data, and Management in a Library Imaging Program.” How can METS be used for digital images with different external schema, in such a way that metadata is integrated?

Although this article is mostly on digital imaging systems and processes – and on the uses of Dublin Core, Content Standard for Digital Geospatial Metadata, and PREMIS with these digital objects – there is a section on the use of METS with PREMIS. The article explains “the Library [of Congress] is integrating broad imaging and metadata standards, work processes, and technical systems with those in other libraries, archives, and museums” (France, F. G., Emery, D. & Toth, M. B., 2010, p. 44). Through collaboration across the cultural heritage community, a range of organizations – including those with large amounts of digital data on cultural artifacts – can meet their advanced digital imaging needs.

The Digital Library Federation publishes the *METS>Metadata Encoding and Transmission Standard: Primer and Reference Manual*, and the latest version was made available as an online PDF in 2010. The 148-page manual provides extensive technical documentation of METS, and it includes examples of the seven METS subsections (Digital Library Federation, 2010). It is an excellent publication for those who need in-depth, technical METS information, and would like an understanding of its complex functions.

In addition to the *D-Lib Magazine* articles, there is thorough information on METS interoperability through the *Interoperability of Metadata for Thematic Research Collections*

report (University of Nebraska–Lincoln, 2012). The report was from a project on which the University of Nebraska, the University of Virginia, Columbia University, and Brown University all collaborated. They studied interoperability issues with METS, MODS, EAD, and TEI using a model based on the Walt Whitman Archive (Bolin, March 2013). The goal of the research team was to create a METS profile for digital thematic research collections. In doing so, they discovered problems with using METS for this purpose. For example, “METS Profiles may become too narrative and thus not machine-actionable” (University of Nebraska–Lincoln, 2012, Conclusions section).

For an introductory profile into some more recent uses of METS, including its use with PREMIS, Lindsey Wagner’s 2011 article “METS: A Survey of Recent Literature and Applications“ provides an excellent summary. The author describes the history and structure of METS, and discusses how METS can be used for federated searching, with administrative metadata, and to archive eJournals (Wagner, 2011).

Elinor Mazé’s 2012 “Metadata: Best Practices for Oral History Access and Preservation,” is a very interesting, well-written, and detailed article that includes information on using METS with digital oral history collections. Mazé is the Senior Editor at Baylor University’s Institute for Oral History. According to Mazé (2012), METS is used for digital libraries, and can be used for collections of digital audio as well as digitized analog sound recordings. In addition to METS, she also discusses other metadata standards, such as Dublin Core, PBCore, MARCXML, EAD, OAI, MODS, PREMIS, and TEI. Since many oral histories are now being digitized from analog tape recordings or born digital through digital recordings, METS can be a useful schema for using metadata with these digital audio files. Mazé points out that the METS schema was applied to digital audio recordings in the Indiana University Digital Library Program Sound Directions

project. She suggests this project could be used in determining best practices for the description and preservation of oral histories.

Finally, “Metadata Developments in Libraries and Other Cultural Heritage Institutions” by Erik Mitchell is a 2013 chapter in *Library Technology Reports*, published by the American Library Association. Although the author only briefly mentions METS, the chapter is one of the few on metadata that focuses on recent interoperability and sustainability issues across different types of cultural heritage institutions. Mitchell (2013) states, “Since 2007, LAM (libraries, archives, and museums) communities have developed new cataloging and archival processing frameworks,” to meet the changing information needs of patrons (p. 5). The chapter provides a model for understanding component parts of metadata systems, and outlines the process for creating and publishing linked data. It is extremely useful in giving a glimpse into the many changes the cultural heritage community will face in the future, as metadata standards and systems are developed to meet user needs.

Conclusion

In the past ten years there has been an increasing awareness of METS as an important metadata schema, which can be useful as a standard for shared data storage and transmission in digital libraries. Recent articles have been written for those in the cultural heritage community who may find METS beneficial. Even so, the uses of METS are still being discovered in terms of best practices and interoperability. One of the main issues with METS is how it can be used with external schemas such as MODS, PREMIS, or Dublin Core.

METS is an excellent metadata schema for use with digital libraries, including those with audio, video, and images as digital objects. Whether these resources are digitized or born digital, METS can be used for the digital files along with a multitude of external schemas, such as

MODS for descriptive metadata or PREMIS for preservation metadata. The Library of Congress provides great resources to help implement METS. With a minimal knowledge of XML, METS is relatively easy to use. As such, METS would be useful in any library, archival, or museum collection that contains digital resources.

References

- Bolin, M. (2013, March). Interoperability: Harvesting. Lecture presented at San José State University, San José, CA
- Bolin, M. (2013, February). MODS, METS, and PREMIS. Lecture presented at San José State University, San José, CA
- Cantara, L. (2005). METS: The metadata encoding and transmission standard. *Cataloging & Classification Quarterly*, 40(3-4), 237-253. DOI:10.1300/J104v40n03_11
- Corporation for National Research Initiatives. (2013). *D-Lib Magazine: The Magazine of Digital Library Research*. Retrieved from <http://www.dlib.org/dlib/alliance-participants.html>
- Cundiff, M. V. (2004). An introduction to the metadata encoding and transmission standard (METS). *Library Hi Tech*, 22(1), 52-64.
- Dappert, A., & Markus, E. (2008). Using METS, PREMIS and MODS for archiving ejournals. *D-Lib Magazine*, 14(9/10). Retrieved from <http://www.dlib.org/dlib/september08/dappert/09dappert.html>
- Digital Library Federation. (2010). *<METS> Metadata Encoding and Transmission Standard: Primer and Reference Manual*. Retrieved from <http://www.LC.gov/standards/mets/METSPrimerRevised.pdf>
- Elings, M. W., & Waibel, G. (2007). Metadata for all: Descriptive standards and metadata sharing across libraries, archives, and museums. *First Monday*, 12(3). Retrieved from <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1628/1543>
- France, F. G., Emery, D. & Toth, M. B. (2010). The convergence of information technology, data, and management in a library imaging program. *Library Quarterly*, 80(1), 33-59.

- Guenther, R., & McCallum, S. (2003). New metadata standards for digital resources: MODS and METS. *Bulletin of the American Society for Information Science and Technology*, 29(2), 12-15. DOI:10.1002/bult.268
- Guenther, R. (2008). Battle of the buzzwords: Flexibility vs. interoperability when implementing PREMIS in METS. *D-Lib Magazine*, 14(7/8). Retrieved from <http://www.dlib.org/dlib/july08/guenther/07guenther.html>
- Library of Congress. (2012). *Metadata Encoding and Transmission Standard (METS) Official Web Site*. Retrieved from <http://www.loc.gov/standards/mets/>
- Library of Congress. (2011). *METS: An Overview and Tutorial*. Retrieved from <http://www.loc.gov/standards/mets/METSOverview.v2.html>
- Library of Congress. (2011). *METS Example Documents*. Retrieved from <http://www.loc.gov/standards/mets/mets-examples.html>
- Library of Congress. (2011). *METS Extenders*. Retrieved from <http://www.loc.gov/standards/mets/mets-extendere.html>
- Library of Congress. (2012). *METS Tools and Utilities*. Retrieved from www.loc.gov/standards/mets/mets-tools.html
- Mazé, E. A. (2012). Metadata: best practices for oral history access and preservation. *Oral History in the Digital Age*. Retrieved from <http://ohda.matrix.msu.edu/2012/06/metadata/>
- Mitchell, E. T. (2013). Metadata Developments in Libraries and Other Cultural Heritage Institutions. *Library Technology Reports*, 49(5), 5-10. DOI:10.5860/ltr
- National Information Standards Organization (NISO). (2004). *Understanding Metadata*. Retrieved from <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

Pearce, J., Pearson, D., Williams, M., & Yeadon, S. (2008). The Australian METS profile: A journey about metadata. *D-Lib Magazine*, 14(3/4). Retrieved from

<http://www.dlib.org/dlib/march08/pearce/03pearce.html>

Riley, J., & Shepherd, K. (2009). A brave new world: Archivists and shareable descriptive metadata. *The American Archivist*, 72(1), 91-112.

University of Nebraska–Lincoln. (2012). *Interoperability of Metadata for Thematic Research Collections*. Retrieved from

http://cdrh.unl.edu/projects/pages/interoperability_metadata.php

Vermaaten, S. (2010). A checklist and a case for documenting PREMIS-METS decisions in a METS profile. *D-Lib Magazine*, 16(9/10). Retrieved from

<http://www.dlib.org/dlib/september10/vermaaten/09vermaaten.html>

Wagner, L. (2011). METS: A survey of recent literature and applications. *Library Philosophy & Practice*, 148-154. Retrieved from

<http://www.webpages.uidaho.edu/~mbolin/lpp2011.htm>