Reimagining METS: An Exploration

Draft Draft Draft for Discussion at Fall 2010 DLF

October 2010

Introduction

Approximately 10 years ago, a group of early digital library innovators and collaborators got together to solve a problem that had been troubling them as they created, exchanged, and retained digital resources over time. As the number, type and complexity of digital resources increased, it became more and more important to find a means for encoding the structure of their digital resources, containing the various metadata associated with the digital resources, and exchanging both the resources and their metadata among repositories. The outcome of the problem-solving collaboration among the nascent Digital Library Federation members was version 1.0 of the Metadata Encoding and Transmission Standard (METS) schema, expressed as XML.

Since 2001, the METS schema has been adopted fairly extensively throughout cultural heritage communities around the world, and currently stands at version 1.9 with only incremental changes occasionally made. While the METS schema has reached a level of maturity and adoption for certain uses in digital library and digital preservation venues, the landscape in which METS is placed is continuing to change. Such changes suggest that it may be time to re-evaluate some key aspects of METS e.g., how it is modeled and expressed.

As the organization responsible for monitoring and adapting the METS schema, and thus continuing its ability to solve the problems for which it was created, the METS Editorial Board (MEB) has begun the process of reimagining METS. As part of that process, this White Paper has been drafted by MEB members and is designed to provide a common level of understanding about METS. The White Paper will discuss contextual issues that might or might not suggest a rethinking of the goals for and functions of a metadata scheme like METS and will be the basis of a discussions in several venues. The MEB is particularly interested in exploring the METS community's perception that a continued need for a METS-like metadata schema still exists given the changing metadata and digital library / archive / repository landscape.

Topics include:

1. A Brief History of METS development and goals
2. New Metadata Technologies and Trends
3. Successful Uses of METS: Encoding, Preservation, OAIS Information Packages, Institutions & Categories
4. METS Implementation Issues & Annoyances
5. Options for Future Directions

Also included with this White Paper are the following Appendices:
Appendix A. References
Appendix B. Background Technical Information
Appendix C. Known METS Implementors, a Partial List
Appendix D. Suggestions for Supplemental Reading

1. Brief History of METS

METS took its impetus from the DLF and National Endowment of the Humanities (NEH) funded Making of America II project that commenced in 1997 with UC Berkeley as the lead institution. This project sought to create a digital object standard that could serve as the cornerstone of an integrated, but distributed digital library. The project's digital object standard was subsequently formalized by Jerome McDonough in the MOA2 XML DTD.

In early applications of the Encoded Archiving Description standard (EAD), implementors wished to provide internal links to individual materials belonging to the collection the EAD described. The MOA2.DTD promised the possibility of solving the problem of how to inventory and organize the numerous digital content files comprising a single digital entity (such as a diary, letter, photo album or scrapbook) and link these content files as well as the entity as a whole with pertinent metadata. Daniel Pitti, the principal developer of the EAD, described this problem as follows:

> "While creating page images was not a problem, it only produced disconnected files. While we could have used clever file names to order these files, we thought this insufficient in the long run from the perspective of access to and managing the "object." What I thought was necessary was an apparatus that would carry descriptive control data as well as information on the logical interrelation of the files."

The MOA2.DTD was in its inception archival - centric, and focused on digitally imaged and transcribed materials. It provided no accommodation for audio and video material. Furthermore, its development preceded all descriptive and administrative metadata XML standards, and thus attempted to define an internal vocabulary for these needs.

Many institutions besides those directly involved in the MOA2.DTD development adopted or experimented with its use. Among these institutions a consensus grew that further development and refinement was needed. In February of 2001, under the sponsorship of DLF, a revision process commenced that led to METS.

Initially METS differed from its predecessor in two main ways. First, building on the key mechanism in the MOA2..DTD that applied structure to content, more sophisticated and varied means for applying structure were developed which allowed METS to be used for different kinds of digital content, such as audio and video resources. Secondly, METS discarded the descriptive and administrative metadata vocabularies defined by the MOA2.DTD since numerous, community defined XML standards for these metadata had started to emerge. Instead, METS provided for the use of these community based standards within specific contexts in a METS document.

METS has gone through many incremental versions since it was first released. Later versions have provided other functionality such as the application of executable and other kinds of behaviors to content files, and the linking of segments of content in a non-hierarchical, hyperlink style fashion. METS is still in its first integral version, however, and its core purposes and mechanisms for accomplishing these purposes have remained unchanged.

2. New Metadata Technologies and Trends

At the time METS and similar digital library standards were developed, metadata development for digital resources and digital repositories often followed an organic, evolutionary process than has subsequently occurred. That is to say, partly because digital libraries and archives were developed from more traditional libraries and archives, the metadata standards for digital resources were often adaptations of already existent element sets with additions deemed necessary or useful for the non-physical aspects of reformatted physical resources. In these cases, it did not seem necessary or useful to start from an abstract data model point of view, as there was little of that being done at the time. Each community started from what they already knew with their inherent assumptions for finding, using and distributing resources that they already knew. For example, libraries assumed that digital content would be searchable in ways compatible with library catalogs and "pathfinders" or bibliographies while archives assumed that a more common path would be finding aids for digital resources within a larger collection or set. Clearly, those assumptions while still being true in many cases, are inadequate for the ever expanding means by which consumers of digital resources find, use and share digital resources. [1]

Since that time, more metadata standards have developed from an abstract data model that attempts to describe a more general approach to a particular problem, and then provide various ways of expressing or binding that data model, e.g., in RDF and/or in XML. At the time METS was developed, XML was the "hot" technology as it extended the proven SGML mark up technology. It was a natural choice for METS given that much of the content described by METS was and is in XML, and much of the metadata collected about objects is in XML or has been easily transformed. [2] See an excerpted selection from the article "XML Fever" for an amusing perspective on XML. [3]

Now, 10 years hence the Semantic Web technologies that were in their infancy and not widely adopted in the digital library community have become more prevalent. Many of the Semantic Web technologies are expressed using the Resource Description Framework (RDF). "RDF is a standard model for data interchange on the Web. RDF has features that facilitate data merging even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed." http://www.w3.org/RDF/. XML and RDF should not be seen as competing but complementary. XML is a syntax and RDF is a data model. Indeed, RDF can be expressed in XML. For further reading on the topic of translating between an RDF based metadata standard and METS, see the ontology devised by Jerome McDonough, et al for OAI-ORE and METS. [4]

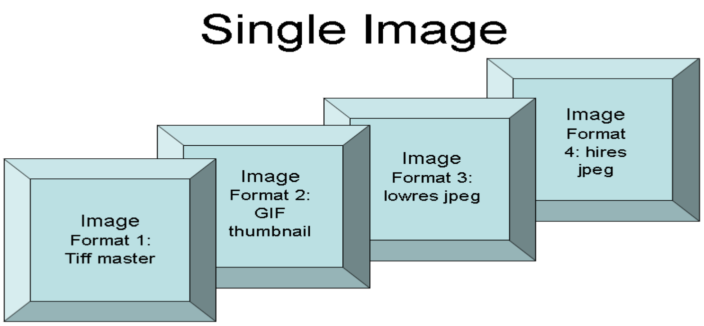
3a. Successful Uses of METS: Encoding

METS is being used successfully to encode the structures of a diverse array of digitized content including textual, pictorial, audio and video materials. These materials may be digitally represented by one or a combination of images in different formats as well as plain and encoded text, digital audio and digital video files. They may be manifested as a single intellectual entity with the physical structure of component files or file fragments to parallel or illustrate the inventory of file formats or content types. Others may be manifested as a single intellectual entity, but describe both the physical structure of the files and the logical structure of the content, or simply the logical content of the intellectual entity. More complex digital resources may include more than one intellectual entity. Textual materials, for example, may be imaged as well as transcribed using TEI. This section of the white paper will briefly describe typical applications of METS structural mechanisms to various kinds of content for various purposes. More complete technical explanations as well as exemplar fragments can be found in the Appendix B of this paper.

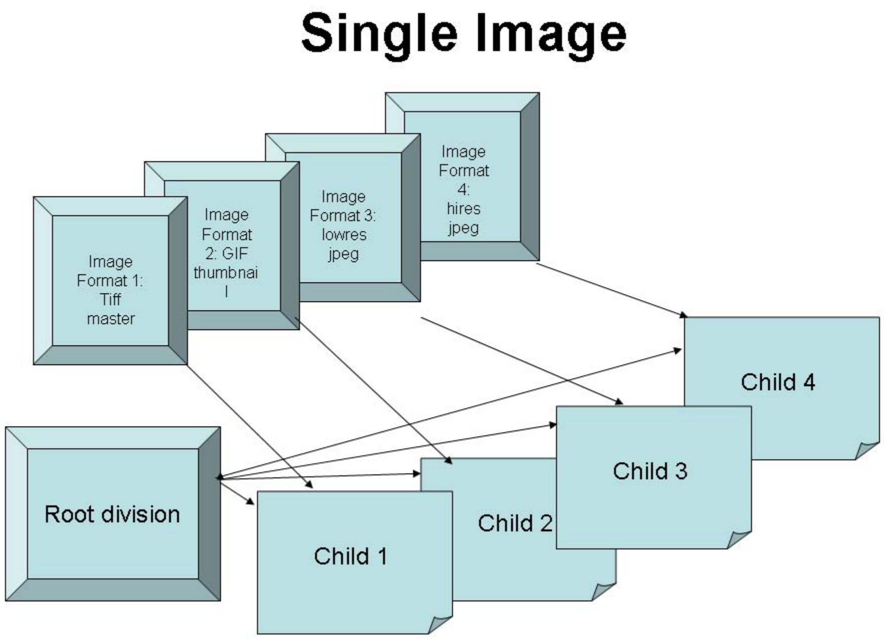
Imaged Content

Both pictorial and textual materials are frequently digitized as images. A digital version of a single pictorial piece, such as a drawing or painting, can illustrate a very basic application of the use of METS.

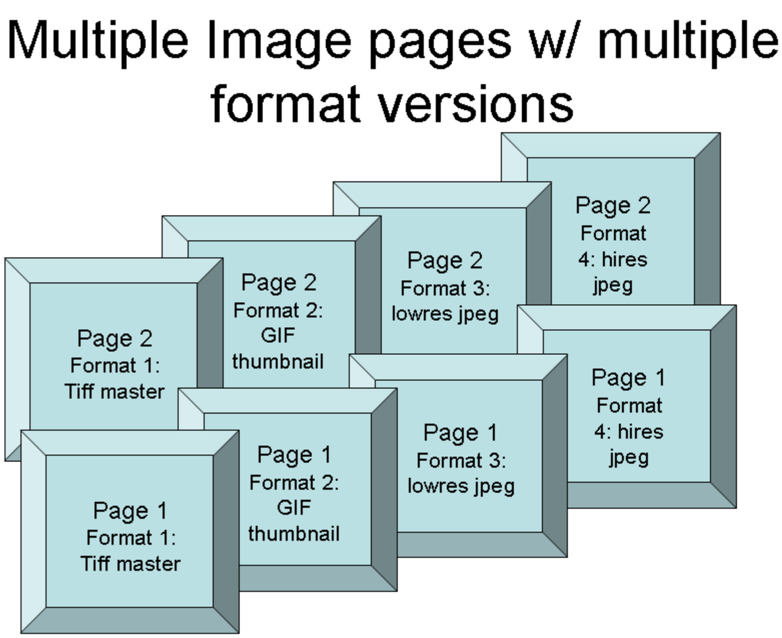
The digital version of a drawing or painting might consist of 3 or 4 images of the same content: a master tiff file, a thumbnail image, and one or two reference or service images.



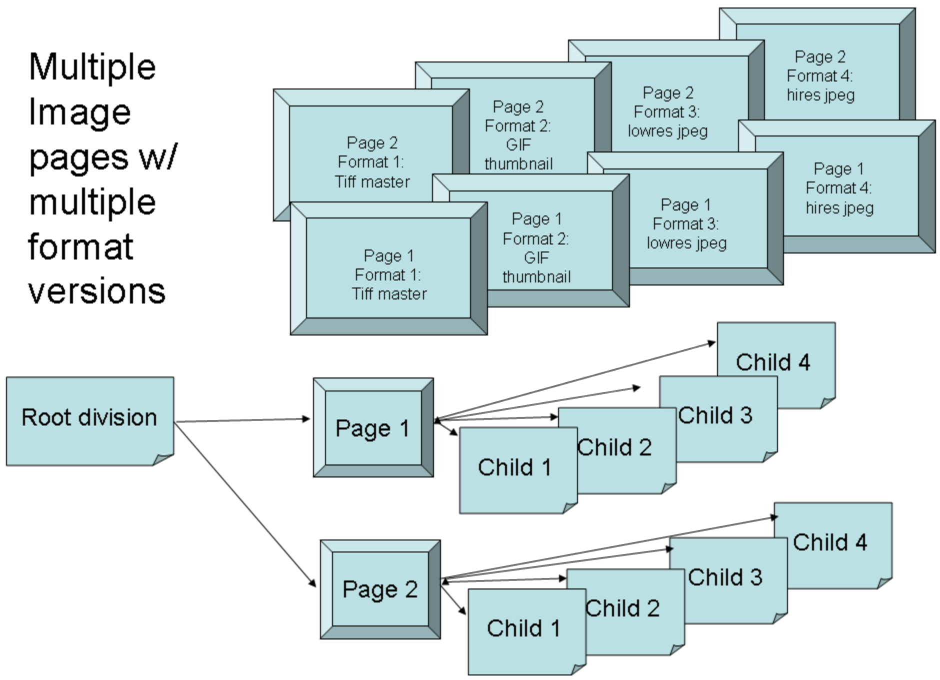
The structural map of the METS document for this case has a single root division which represents the drawing. This root division has four children, each of which points to an integral content file, i.e., one of the four format versions of the drawing. See Example 1 in the Appendix B.



As with single pictorial materials, images frequently comprise the digital version of pictorial materials with more than one image, (such as a photo album) and textual materials of more than one page. Such "compound" materials involve more structural complexity, and a digital version is likely to require multiple imaged versions of each of multiple pages in order to be able to navigate and display the images in order, for instance, or to associate descriptive or administrative metadata with each of the pages.



To recreate the experience of the analog source, structural mechanisms must be able to specify the proper relationship both between the multiple imaged versions of each page, as well as between the various sets of images for all of the pages. Often a purely physical type METS structural map will be used for this purpose in which the root <div> represents the photo album as a whole. This root <div< is analysed into a sequence of child divisions representing the pages of the scrapbook. Each of these child divisions then has 4 <fptr> children, one for each of the available imaged manifestations of the page. See Example 2 in the Appendix B.



If resources are available, some METS implementers choose to use the METS structural map to provide a logical superstructure above the physical structure--or the level at which the imaging has been done. For example, Indiana uses the METS <structMap> to analyze a digitized journal first into a sequence of logical units (front matter followed by a sequence of articles). Each logical unit is then further analyzed into a sequence of physical pages with which the image manifestations are associated. See Example 3 in the Appendix B.

Mixed Content: Image + TEI, PDF, Word, etc

As previously noted, in the case of digitized textual materials, often more than one type or format of the content needs to be made available. For example, in addition to the imaged versions of the individual pages, the text may be available as a TEI transcription, as plain OCR'd text or in MS Word or other word processor format. Sometimes downloadable PDF versions of individual parts or all of a digitized document may be available. The granularity of the alternate digital formats may be different from the imaged formats. While imaging is typically done at the page level, a TEI transcription would typically represent an entire document. PDF versions often represent an entire document; but might also be made for individual logical units such as the article level of a journal, for example, or physical units of the page. METS implementers handle these various cases in several typical ways.

Similar Level of Granularity: Where the granularity of the different available digital formats is the same, then available manifestations can simply be referenced by sibling <fptr> elements under a common parent. See Example 4 in the Appendix B where both OCR and imaging has been performed at the page level, and the page level OCR and image files are referenced by sibling <fptr> elements under <div> elements representing the individual pages.

Differing Levels of Granularity: Implementers typically use one of three encoding strategies for cases where the granularity of the available digital content formats differs.

- Separating the available formats into discrete, sibling branches of the structural map hierarchy. See Example 5 in the Appendix B.
- Associating the different types of content files with the appropriate relative level of the document hierarchy such as representing the articles in a journal as PDFs, and also including the individual images for each page of the article. See Example 6 in the Appendix B.
- Associating a TEI transcription and imaged versions of textual materials. See Examples 7 and 8 in the Appendix B.

NOTE: Alternative approaches to encoding textual materials that are scanned at the page level, but made available at the article level can be found in the Appendix B entitled "METS Textual Capabilities Extended with ALTO", and associated Examples 9 and 10.

METS with A/V Materials

Implementers using METS with A/V materials frequently draw on the capabilities of the METS <area> element to divide an integral audio or video file into logical segments; and the <area> element supports numerous standards for marking the beginning and ending of such logical A/V segments. Examples of various approaches for marking the logical A/V segments can be found in the Appendix B at:

- Example 11 where the Library of Congress has obviated the need to use the <area> element by creating segmental derivatives from an integral master, so that each segment is manifested by an integral derivative file.
- Examples 12, and 13 show how Indiana University and other institutions apply structure to an integral audio file by specifying time code values with various attributes of an area element associated with each logical segment.
- Example 14 shows how Harvard University chooses to have the structMap defer to ADL files for recording internal structure of audio files.

NOTE: For other examples of how the <area> element can be used to apply structure to an archive file, see Example 15 in the Appendix B in "Using BYTE Offsets".

Powerful but lesser used METS capabilities for encoding structure

METS has capabilities for encoding digital resources that appear to be used less frequently, but are still helpful for those institutions which use them. See the Appendix B for information about "Multiple Files Manifesting a Structural Division" in which various means are discussed for specifying the structure of multiple files of parts of files that must be played or displayed in parallel or in sequence. Examples 16 – 18 illustrate use of the sequence <seq> and parallel <par> elements. In addition, the "Non-hierarchical Linking" section in the Background discuss the use of the <structLink> element that allows internal linking between divisions between structural maps or within the same structural map. Examples 19 and 20 illustrate various approaches to use of this element.

METS and EAD

METS is frequently used in conjunction with EAD finding aids. The EAD describes a collection as a whole, as well as the individual entities comprising a collection. When an individual intellectual entity that is represented by a finding aid is available in digital format, frequently the digital version will be encoded as a METS object, which is then referenced from a <dao> element associated with the corresponding item in the EAD item list. METS provides for referencing the finding aid or finding aids in which the intellectual entity represented by the mets object participates in a dmdRef element; it can also, via an XPTR attribute, specify the specific element in the item list that corresponds to it. There are several tools (such as Archivist Toolkit) that are geared towards creating METS objects in conjunction with EADs.

3b. Successful Uses of METS: Preservation

Metadata is considered vital to keeping digital content usable over time. METS is widely used in digital preservation repositories and projects as a mechanism for aggregating, coordinating, and managing the metadata associated with a digital object and its component files.

Aggregating metadata and content for an object

METS supports several options for compiling content and metadata associated with an object: referencing through links, encapsulating as in-line XML, encapsulating as base-64 encoded data, and use of the METS pointer or <mptr> element. METS defines segments for descriptive, administrative (technical, source, rights, digital provenance), and structural metadata, a file manifest and associated behaviors. Repositories using METS to carry the metadata needed for preservation can then manage that METS file as a primary object, giving it the same degree of secure storage and management as the content files themselves. Using METS in this way also supports object portability.

Through the <mptr> element, METS provides for aggregating and applying a structure to multiple external METS objects. The ECHO Dep project, for example. uses a higher level METS object to aggregate and order METS objects representing web captures made on different days for preservation purposes. See Example 21 in the Appendix B. Columbia uses this feature to draw together disparate mets files that represent the individual audio source files as well as the rendered audio and service files for a single, intellectual audio entity. See Example 22 in the Appendix B.

The general, repeatable "bucket" elements -- for descriptive, technical, rights, source and digital provenance metadata -- can contain specific metadata defined by other schemes. This flexibility makes METS usable by a wide range of communities that require metadata (commonly descriptive and technical metadata) tailored to their domain. Of particular strength and growing use is the combination of METS with the PREMIS standard.

METS and PREMIS

PREMIS, the data dictionary and schema for preservation metadata, can be used in conjunction with METS. Using PREMIS to augment METS strengthens the preservation function by supporting a consistent baseline of technical, administrative and provenance metadata.

METS, PREMIS, MIX and other metadata standards relevant to preservation of digital content were typically developed to stand alone -- no assumptions were made that particular standards would always be used together. For this reason, standards overlap in basic information such as checksum, size and format. A joint PREMIS/METS working group has developed best practice guidelines for using PREMIS within a METS context, to improve consistency in the way metadata is populated across repositories and, in consequence, improve preservability of the digital objects. [5]

Preservation Management: Tracking states of an object

METS has also been used in a so-called "hub and spoke" model, where individual states of the object are captured as it is edited over time, each state represented by separate METS file. These snapshots are themselves the content files of another "hub" METS document that represents the object throughout its history. The ECHO DEPository project, funded by NDIIPP, developed the Hub and Spoke (HandS) tool suite "to help curators of digital objects manage content in multiple repository systems while preserving valuable preservation metadata. Implementing METS and PREMIS, HandS provides a standards-based method for packaging content that allows digital objects to be moved between repositories more easily while supporting the collection of technical and provenance information crucial for long-term preservation."[6]

3c: Successful Uses of METS: OAIS Information Packages

From its inception, METS was intended to support both "management of digital objects within a repository and the exchange of such objects between repositories." [7] A METS document, encapsulating or linking to content and metadata for a digital object as a whole or any of its components, can serve as any of the "packages" defined in the OAIS Reference Model: a submission information package (SIP), an archival information package (AIP) and a dissemination information package (DIP).[8] While concrete evidence has not yet been collected, it appears that most common uses for METS in terms of OAIS information packages are for SIPs and DIPs, particularly for repository systems such as Dspace and Fedora as well as for some commercial vendors including ExLibris (an integrated library system vendor) and ccs:docWorks, a commercial digital services vendor.

3d: Successful Uses of METS: Institutions and Categories

The most common uses of METS have been by cultural heritage institutions including national and state libraries and agencies, college & university libraries, public libraries, and consortial / shared digital libraries as evidenced by the partial list of institutional names included in Appendix C. As more open source and proprietary tools have become available for METS creation, navigation and use, the number and types of institutions have expanded. [9]

4. METS Implementation Issues & Annoyances

The topics included in this section of the White Paper represent areas of some discussion and concern by either METS implementors, MEB members or both. They serve as background for the next two section in which questions, problems and possible solutions are considered briefly. These two sections are the basis upon which the MEB wishes to engage the METS community in discussion at the Fall DLF Forum and other venues as feasible and desirable.

1. **Archiving of web resources**

For the harvesting of Web pages special tools called crawlers are used. These tools scan web pages for links to follow and save the pages on its way. The most used crawler for web archiving is the open source software Heritrix. [10] The results of each crawler session with Heritrix are stored in a container format called ARC or WARC. [11] WARC is the updated Version of ARC and an ISO standard. It is widely adopted by web archives and used by several tools like the Wayback Machine. [12]

Web archiving is often done by institutions which are also maintaining digital archiving systems. As METS is widely used for archiving packages in these OAIS compliant systems, there is a need to adopt METS for the results of web harvesting. WARC is a binary container format for files and metadata and METS is a XML container format for files and metadata. The challenge to bring these two concepts together could be solved in different ways:

- The WARC files could be treated as content files. In this case the metadata in METS would be just about the WARC files itself and not about the included files in the WARC files. The advantage would be an easy integration in existing archiving workflows. The disadvantage is the lack of information for preservation strategies. In most archival systems PREMIS is used within METS to describe the objects and to be able to execute preservation actions like migration and emulation. But without the information about the included file formats in the WARC containers it will be nearly impossible to do migration on the level of the METS/PREMIS description.
- The WARC container could be unpacked and the separated files could be handled in METS just like any other files. The WARC container includes also information on the crawls, so this data would have to be stored in new files. Another problem would be a proper access tool. The WARC files are optimized for special tools like the Wayback Machine that automatically change the absolute links in the web pages to relative ones and do other adjustments to allow a proper archival access to old web pages. All this handling would have to be done by the archival system.
- The WARC files and the included files are listed and described by metadata in METS. This would give the ideal basis for further preservation actions. The main problem with this approach is the amount of data. A typical WARC file could contain thousands of files. A PREMIS record would have to be created for every file and put in the METS file. In order to create the PREMIS metadata file characterization and identification tools like DROID and JHOVE are used. These tools are not (yet) optimized to be used for thousands of files.

When web archiving started most of the institutions just collected the ARCs separately from there archival systems. But preservation issues are also relevant for web archives and this leads to the integration of the existing ARC/WARC files in the systems for digital preservation. METS and PREMIS are used in these systems and the institutions try to find ways to integrate the web files in existing profiles. Although it would make sense for preservation planning, for pragmatic reasons it is not possible to generate metadata for every file within the WARC files.

A recent example of a way to deal with the need for preservation and the limitations of current tools is the METS profile for web archiving by the British Library:
http://www.ifs.tuwien.ac.at/dp/ipres2010/papers/enders-70.pdf

2. Requiring the organization of digital resource associated metadata into required "buckets"

The metadata taxonomy of METS—from the broad dichotomy between administrative and descriptive metadata as represented by the <amdSec> and <dmdSec> to the specific buckets, e.g., <techMD> and <digiprovMD>—has met with mixed approval from METS' implementors points of view. It is not always clear in what bucket to put the metadata; hence there is some discontinuity among implementors. One of the purpose the prescribed buckets serve is to get implementors to think about different categories that might be important; nevertheless, questions have been raised as to whether this and other reasons for proscribing the metadata categorization are sufficient .

3. Using METS as a digital resource exchange format

Use of METS for exchange has not been overly successful to date. Some writers (Maslov, et al [13]) have described METS as less an exchange format and more of a packaging format because of the flexibility inherent in it. From an exchange point of view there are too many possibilities for declaring where and how both content and metadata can be found. Also, the definitions of what is needed for exchange are not clearly defined, thus making the exchange packets represented by METS documents moving targets.

Other writers have corroborated the above comments to some extent in comparing the approach of METS to OAI-ORE for making digital resources available over the web (McDonough [14]) In his article, McDonough states that the requirements that OAI-ORE places upon those wishing to make digital resources available are comparatively simple. From McDonough's analysis, it is relatively simple to make METS documents that are not too complex accessible to OAI-ORE resource maps by following recommended practice for METS creation. Ironically, the very fact that the METS schema allows in depth descriptions of complex resources in a number of different ways makes meeting the simplicity of OAI-ORE's requirements more difficult. From an OAI-ORE point of view, it can it be difficult to untangle the relationships among files or file fragments in order to describe a unit of exchange ("aggregation") described by a resource map, as one may need to look within either or both the mets:fileSec for fileGrp/file and the mets:structMap for div/fptr and/or div/fptr/area.

Another factor has been cited both by McDonough and noted in the IEEE RAMLET [15] work in which various content packaging standards such as METS, IMS-CP, MPEG DIDL and OAI-ORE are mapped to a derived ontology. That is, the fact that much of the descriptive metadata about a digital resource useful for citing in OAI-ORE resource maps and other citation mechanisms is buried in attributes within elements within METS, such as file size, and checksum values and algorithms. In addition, reliance upon the XML ID/IDREF/IDREFS mechanism for internal referencing is somewhat inconsistently applied in METS, and not necessarily feasible for creation of required (from an OAI-ORE point of view) or at least advisable (from any later citation of a resource) URIs. Use of the xLink standard within METS does mitigate the difficulty of creating URIs to some extent, if used with full intention beforehand of making digital resources citeable and web accessible.

By contrast, recent work done by the Florida Center for Library Automation via the TIPR Project[16] illustrates how METS and PREMIS have been successfully used for the exchange of digital resources among three different libraries, with pre-planning. SeeDiscuss use of METS as the basis for TIPR and for the FCLA work as the SIP for the Florida Digital Library Archive.

Use of the METS profile schema would also greatly facilitate the clarity of requirements associated with exchange of a given application of METS to a class of digital resource or to a given institutional METS implementation.

4. Fedora and METS

In 2000/2001, as initial work on the Fedora repository software was underway at the University of Virginia, the development team needed a markup language capable of expressing the full nature of a digital object. Their goal was to use and existing standard, and to adopt one that seemed to have some standing in the digital library / digital repository communities. Their choice was METS.

The METS standard provided a means to bind together datastreams of various types (content datastreams, metadata datastreams) and to provide a mapping between these items. It did not, however, provide a mechanism for effectively linking a digital object to the mechanisms needed to access and use the content. Since versions 1 and 2 of the Fedora system explicitly included disseminators in the metadata for each object (version 3 introduced the Content Model Architecture, where disseminators are first class objects themselves and content objects can subscribe to appropriate disseminations) there was a need to extend the existing schema to accomplish this task. To this end, the behaviorSec element group was added to the schema in 2001.

Despite the initial decision to employ METS as the primary object modeling language, the development team opted to create a new language (FOXML) and employ it as the primary language. This decision was made for several reasons:

- Simplicity – user feedback called for a conceptually easy mapping of the Fedora concepts to an XML format. Users wanted an obvious sense of how to create Fedora ingest files, especially those who are not familiar with formats such as METS.
- Optimization and performance – It was felt that the schema was too strongly typed and was not abstract enough. Additionally, initial testing showed that the high level of indirection led to costly processing at scale.
- Flexibility – establishing FOXML as the internal storage format for Fedora objects enables easier evolution of functionality in the Fedora repository, without requiring ongoing extensions to the METS schema.

Despite the establishment of FOXML as the primary expression of the Fedora object model, there are instances where METS and Fedora may continue to operate:

- Fedora continues to support METS for ingestion of objects. It does so through a Fedora specific extension of the 1.1 version of the METS schema couple with very tightly controlled rules.
- Transport and exchange of objects between repositories with different repository architectures.

- As a structural metadata datastream, especially for inclusion and serialization of complex objects such as page-imaged books.

5. Use of profiles to achieve "interoperability"

Shortly after the first versions of the METS schema were developed, the MEB developed a profile schema to help implementors document how they were using METS and for what kinds of digital content. Ostensibly to help facilitate "interoperability" among METS users, the profile schema, while expressed in XML, was intended to be human readable rather than machine readable. The MEB also devised a registration process for METS profiles to help manage versions and publicize their existence. Over time, the use of METS profiles has been less successful than the MEB wished, and efforts have been made to improve it, resulting in version 2.0 of the METS profile schema to be released in early 2010. Some thoughts about the utility of the METS profiling approach follow.

> "However, METS is primarily a packaging format and not an object exchange protocol. Without an explicitly defined profile, METS lacks the specificity necessary to force a consistent interpretation of the encoding. While this makes METS very versatile for encoding different object types, in the context of repository interoperability this results in a brittle solution." [17]

> "The digital library community seems to face a dilemma at this point. Through its pursuit of design goals of flexibility, extensibility, modularity and abstraction, and its promulgation of those goals as common practice through its implementation of XML metadata standards, it has managed to substantially impede progress towards another commonly held goal, interoperability of digital library content across a range of systems.
>
> How then, should the community respond?
>
> One possible response to this situation would be to say that perhaps our community cares less about interoperability than we thought...it may be that interoperability is in fact a lower priority for the digital library community than it likes to believe, and the adoption of metadata standards that impede interoperability is merely a reflection of that underlying reality, and not a major problem to resolve. There is at least some reason to suspect this may be the case. Most research libraries have a clearly defined local clientele...libraries' primary responsibility will always be to their local communities....many libraries may decide that interoperability, while desirable, is a goal which may have to wait." [18]

Profiles are monolithic / hard to pick and choose features. Because there are so many choices inherent in the practice of METS description, a profile tends to be idiosyncratic and tied to a specific system implementation. There is no way easy way to say "I'll take the structMap from this profile, but I need to use my own descriptive metadata standard."

On the Practice of Profiling METS Documents, see https://www.socialtext.net/m/page/mim-2006/practice_of_profiling

Profiles at best allow one to document a community of practice for a defined purpose, but because METS supports such diversity of valid practice profiles tend to hinder rather than enable true interoperability. Tools written for METS tend to only work with a very specific set of profiles.

Further, Profiles merely present documentation for subclasses of METS documents. Even if concerned with targeting a particular Profile, one still needs to interpret the Profile and develop processes to actually produce, consume, or otherwise process conformant METS instances. Acknowledging that human readable documentation is enormously useful, it may be questioned how much more useful a METS Profile is over say a well documented Schematron schema, or subset XML schema. Indeed, it might be a valuable METS Board activity to produce and maintain a few "endorsed" generic Profiles (e.g., image collection, text with images, etc...) with accompanying schemas as much as possible expressing the Profile's constraints.

See the Supplemential Reading section for more information on this topic including an article on "Aligning METS with the OAI-ORE Data Model" by Jerome P. McDonough.

6. Options for Future Directions – some thoughts for discussion

- What might a re-imagined METS look like? The current design of METS is setup for generica packaging with specific meanings either unspecified or slightly more specified via profile and local system design. This approach is highly flexible but perhaps doesn't really facilitate interoperability? Would an alternative (SemWeb inspired) approach be providing formally defined relationships that can be used in any needed combination. Is there anything in the middle? Which direction is most useful for METS?
- Which of these functions should a new or formally expressed METS data model take on?
    - Packaging function - multiple files plus metadata needed for interpretation and use of them (including historical/preservation actions information)
    - Structure internal to a "resource"
    - Describing relationships between closely tied "resources" (e.g data sets and articles published on them a la ORE)
    - Behaviors - how specific files, resources should be used/displayed/etc.
    - Should METS be optimized for exchange (interoperability) or for local storage / use?
- Should standardized vocabularies for USE and TYPE attributes be set up and register with http://www.schemaweb.info/?
- What's the right (or most useful) relationship to OAI-ORE, BagIt, FOXML, etc?
- Going forward, should the Board be more concerned with *application* of METS rather than its *design*? Perhaps METS is OK as it is and that what people really need is assistance in the practicalities of creating and processing it.
- Are profiling mechanisms necessary? What about machine-actionable profiling mechanisms?http://www.schemaweb.info/

Appendix A: References

---

[1] For more discussion on this topic see some notes on METS as a theory of digital object description http://tingletech.tumblr.com/post/1186443159/.

[2] METS is slightly biased towards w3c xml schema (XSD) in that this is the XML schema technology used to validate METS files and any attached metadata that has an associated schema. It is not completely clear how to embed schema information for embedded XML conforming non-w3c xsd schema such as Schematron or RELAXNG. Some have suggested that NVDL (Namespace-based Validation Dispatching Language: http://www.nvdl.org/) can be used for this. The MEB could provide guidelines on using NVDL with METS.

[3] "XML Fever" by Erik Wilde and Robert J. Glushko, http://dret.net/netdret/docs/wilde-cacm2008-xml-fever.html

> "XML fever can be acquired in many different ways, but the most prevalent way is to be infected by the idea that XML enables almost magical universal interoperability of information producers and consumers. XML fevers can be classified as basic, intermediate, and advanced:
>
> - *"Basic strains* infect XML neophytes, but most of them recover quickly. It can be disappointing to discover that the landscape of XML technologies is not as simple as expected, and that working with the associated tools requires some getting used to, but most people develop some immunity to the XML hype and quickly begin to do useful work with it.
> - *"Intermediate strains* of XML fever are contracted when XML users move beyond simple applications involving structured information and encounter models of data, documents, or processes. A recurring symptom in these varieties of XML fever is mild paralysis brought on by having to select a schema language to encode a model, trying to choose among the bewildering number of features in some languages, or trying to round-trip a model between different environments.
> - *"Advanced strains* of XML fever often take hold after exposure to the proliferation of more complex and esoteric XML-based technologies layered on top of it. These advanced diseases are harder to catch, but they are also harder to remedy because people who have caught these advanced strains tend to congregate with others with the same diseases and they are continually reinfecting each other."

[4] "Preserving Virtual Worlds Final Report; Metadata & Packaging Recommendations"¸ McDonough, Jerome P. et. al., September 2010, http://hdl.handle.net/2142/17097, pgs 102-104.

[5] Guidelines for using PREMIS with METS for exchange Revised September 17, 2008. http://www.loc.gov/standards/premis/guidelines-premismets.pdf

[6] UIUC Echodep Hub and Spoke Framework Tool S. Vuite. http://dli.grainger.uiuc.edu/echodep/hands/ Viewed Oct 4, 2010.

[7] Metadata Encoding And Transmission Standard: Primer And Reference Manual. Version 1.6 Revised. Digital Library Federation, 2010. http://www.loc.gov/standards/mets/METSPrimerRevised.pdf.

[8] Reference Model for an Open Archival Information System (OAIS). Blue Book. Consultative Committee for Space Data Systems, January 2002. http://public.ccsds.org/publications/archive/650x0b1.pdf

[9] More information about METS tools can be found on the METS website at http://www.loc.gov/standards/mets/mets-tools.html.

[10] http://crawler.archive.org/

[11] http://www.digitalpreservation.gov/formats/fdd/fdd000236.shtml

[12] http://archive-access.sourceforge.net/projects/wayback/

[13] Adding OAI-ORE Support to Repository Platform, Alexey Maslov et al, http://journals.tdl.org/jodi/article/download/749/640 [PDF]

[14] Aligning METS with the OAI-ORE Data Model, McDonough, Jerome P. 2009-06 http://hdl.handle.net/2142/10744

[15] IEEE Learning Technology & Standards Committee's Resource Aggregation Models for Learning, Education and Training, https://mentor.ieee.org/ramlet/bp/StartPage.

[16] Towards Interoperable Preservation Repositories, a project funded by a National Leadership Grant from the Institute of Museum and Library Services. See http://wiki.fcla.edu:8000/TIPR

[17] http://journals.tdl.org/jodi/article/download/749/640 page 3 Alexey Maslov et. al.

[18] McDonough, Jerome. "Structural Metadata and the Social Limitation of Interoperability: A Sociotechnical View of XML and Digital Library Standards Development." Presented at Balisage: The Markup Conference 2008, Montréal, Canada, August 12 - 15, 2008. In *Proceedings of Balisage: The Markup Conference 2008*. Balisage Series on Markup Technologies, vol. 1 (2008). doi:10.4242/BalisageVol1.McDonough01.

Appendix B: Background Technical Information

1. Technical Explanation of METS Encoding mechanisms:

METS provides for applying one or more hierarchical structures to digital content through the repeatable <structMap> element. Each <structMap> element can contain one and only one root <div> element, which represents the content of an intellectual entity as a whole. The root <div> element can contain a (potentially) infinitely nested set of <div> or division elements under it. Each <div> element may serve a completely structural function, and not have any content directly associated with it. However any <div> can also be associated with content that manifests the structural division by means of one or more child <fptr> elements. Each <fptr> element represents a complete manifestation of its parent <div> element. Sibling <fptr> elements under a common parent <div> represent alternate manifestations of the content represented by the structural division. (For example, each <fptr> under a <div> may point to a different resolution image of the same content). An <fptr> element can point directly to an integral content file (as represented by a <file> element in the file inventory or <fileSec>). In this case, the integral content file would completely manifest the content of the structural segment represented by the <div>. An <fptr> may also point just to part of an integral content file via a child <area> element, if just part of a content file is needed fully to manifest a structural division. In some cases multiple files or parts of files played or displayed in sequence or in parallel to manifest a structural division. In this case an <fptr> element would contain a <par> or <seq> element, which in turn would aggregate pointers to the files or parts of files needed to manifest a structural division. In rare, and especially complex cases, a <par> element may contain a <seq> element and a <seq> element may contain a <par> element.

2. Technical Explanations of Encoding Examples from Section 3:
The examples below have been streamlined and adjusted to highlight the pertinent structural mechanisms.

Imaged Content:

Both pictorial and textual materials are frequently digitized as images. (In addition, textual materials may be represented by structured or plain text files, or in proprietary textual formats such as Word or potentially composite formats such PDFs).

A digital version of a simple pictorial piece, such as a drawing or painting, can illustrate a very basic application of the mets <structMap>. The digital version of a drawing or painting might consist of 3 or 4 images of the same content: a master tiff file, a thumbnail image, and one or two reference or service images. A METS fragment for the digital version of a drawing owned by the The Bancroft Library provides a typical encoding for this case (Example 1). The digital version of the drawing includes 4 content files: a tiff master, a gif thumbnail, a low resolution jpeg reference image and a high resolution jpeg reference image. The <structMap> for this case has a single <div>--the root <div>--which represents the drawing. This root <div> has four child <fptr> elements, each of which points to an integral content file: one of the four imaged versions of the drawing.

As with unitary pictorial materials, images frequently comprise the digital version (or a digital version) of compound pictorial materials (such as a photo album) and textual materials. Often, however, such materials involve more structural complexity, and a digital version is likely to require multiple imaged versions of each of multiple pages. To recreate the experience of the analog source, structural mechanisms must be able to specify the proper relationship both between the multiple imaged versions of each page, as well as between the various sets of images for all of the pages. Often a purely physical type METS <structMap> will be used for this purpose as in Example 2 based on the METS encoding for a photo album held by Brown University. Here the structMap contains a root <div> element that represents the photo album as a whole. This root <div> is analysed into a sequence of <div> elements representing the pages of the scrapbook. Each of these <div> elements has 4 child <fptr> elements, one for each of the available imaged manifestations of the page.

Resources available, some METS implementers choose to use the METS <structMap> to provide a logical superstructure above the physical structure--or the level at which the imaging has been done. For example, Indiana uses the METS <structMap> to analyze a digitized journal first into a sequence of logical units (front matter followed by a sequence of articles). Each logical unit is then further analyzed into a sequence of physical pages with which the image manifestations are associated. (Example 3).

Mixed Content: Image + TEI, PDF, Word, etc

In the case of digitized textual materials, often more than one type or format of the content is available. In addition to the imaged versions of the individual pages, the text may be available as a TEI transcription, as plain OCR'd text or in Word or other word processor format. Sometimes downloadable PDF versions of individual parts or all of a digitized document may be available. The granularity of the alternate digital formats may be different from the imaged formats. Imaging is typically done at the page level; a TEI transcription would typically represent an entire document. PDF versions might represent an entire document; but might might be made at the level of individual logical units (at the article level in the case of a journal, for example), or physical units (the page). There are several typical ways that METS implementers handle these various cases.

Where the granularity of the different available digital formats is the same, then available manifestations can simply be referenced by sibling <fptr> elements under a common parent. This is the case in Example 4 where both OCR and imaging has been performed at the page level, and the page level OCR and image files are referenced by sibling <fptr> elements under <div> elements representing the individual pages.

Implementers typically use one of three encoding strategies for cases where the granularity of the available digital content formats differs.

Some implementers separate out the available formats out into two separate, sibling branches of the <structMap> hierarchy. Brown University has both PDF and page image versions of its digitized journals, and chooses to segregate these in separate branches of a structMap. (Example 5)

Other implementers prefer a more integrated approach, and simply associate the different types of content files with the appropriate relative level of the document hierarchy. For example, Indiana makes PDF content files for the individual articles in a digitized journal available. It associates these via child <fptr> elements with the level of the mets <structMap> hierarchy that represent the individual articles. The <div> elements representing the article level of the hierarchy are also analyzed into sets of child <div> elements that represent the individual pages comprising the articles. And it's these <div> elements that are associated with the page images. (Example 6).

METS implementers who make both TEI transcription and imaged versions of textual materials frequently use this same approach. For example, in the case of a letter that has been both imaged and transcribed, UCLA has associated the TEI file, which represents the entire letter, with the root <div> element, and the page images with <div> elements representing the individual pages immediately below the root <div>. (Example 7).

Other METS implementers dealing with textual materials represented both by a TEI transcription and page images choose to reconcile the difference in granularity between the structured text and imaged versions of the textual entity. These implementers use the METS <area> to isolate the specific elements within the <body> of the TEI transcription that correspond to the contents of the individual page images. This can be seen in Example 8, drawn from a registered profile submitted by the Whitman archive. Here, each side of each page of the manuscript of Whitman's "Italian Music in the Dakotas" has been imaged separately. The <body> of the TEI transcription uses the <pb> element, each identified by an id value, to mark the page boundaries within the transcription. Within the METS <structMap>, each <div> element that represents the individual side of a manuscript page, includes <fptr> elements that point to each of the imaged manifestations of the side. It also includes as a sibling to these <fptr> elements, an <fptr> element with a child <area> element identifies not just the integral TEI content file, but by means of IDREFS, the specific specific section of the of the TEI corresponding to content of the page images.

METS Textual Capabilities Extended with ALTO

The METS <area> element can also be used to reconcile discrepancies between a desired logical structure, and the way that an intellectual entity has been imaged, or (in the case of A/V materials) captured in an audio or video file. The A/V case will be covered below. With textual materials, such a discrepancy is common with newspapers, where the imaging is done at the page level, but an implementer might want to allow access at the article level. The METS <area> element does provide, via HTML4 style SHAPE and COORDS attributes, for isolating the specific sections of individual images that correspond to a particular logical segment. In practice, this feature appears to be seldom used, probably because of the resources involved. Instead, many implementers have chosen to extend the capabilities of the METS <structMap> with ALTO, an XML standard initially developed by CCS in Germany, but now maintained by an independent Editorial Board and distributed by the Library of Congress. This solution is attractive because its implementation can be largely automated (via docWorks, developed by CCS). In addition, many digitization service providers (generally using docWorks) can provide a coordinated ALTO and image content files along with METS. The National Digital Newspaper Project requires the use of ALTO with METS of its participants. Some of these participants, such as the California Digital Newspaper Collection, are using ALTO in conjunction with detailed logical METS <structMap> elements to provide article level search and display capabilities.

ALTO, which can be considered a kind of "super" OCR file. Typically an ALTO file would capture the physical structure of a page of the analog source. It does so through a hierarchy of physical components comprising the page that go down to the individual word level. At each level the physical location of the corresponding component of the source page is specified by means of coordinates. If the resolution of an image that manifests the source page is known, then the physical coordinates can be converted to image coordinates. For article level access, a logical METS <structMap> that analyzes a newspaper edition into articles is cross-referenced to an ALTO file or files by means of IDREFS that identify the elements in the ALTO file that represent the components making up the articles. A complete description of ALTO and its capabilities is impossible in this context. However, an excerpt of the <structMap> of a digitized newspaper from the California Digital Newspaper Collection (Example 9), and one of it's associated ALTO files (Example 10) can give the flavor of it's capabilities.

METS with A/V Materials

Implementers using METS with A/V materials frequently draw on the capabilities of the METS <area> element to divide an integral audio or video file into logical segments; and the <area> element supports numerous standards for marking the beginning and ending of such logical A/V segments. LC has obviated the need to use the <area> element in conjunction with at least some of the video offerings on its Encyclopedia of the Performing Arts by creating segmental derivatives from an integral master, so that each segment is manifested by an integral derivative file. (Example 11). Indiana and other institutions prefer to apply structure to an integral audio file by specifying time code values (TCF based in the case of Indiana) in the BEGIN and END attributes of an area element associated with each logical segment. (Example 12). An time based alternative is used by LC in encoding CDs that employs the BEGIN element in conjunction with the EXTENT element to isolate an audio segment. (Example 13).Indiana and Harvard participated in a joint Sound Directions project in an attempt to establish best practices, including the use of METS, in conjunction with audio content for preservation and access purposes, but in fact came to rather different conclusions and practices. In terms of its use of the structMap for audio puposes, Harvard prefers to have the structMap defer to ADL files for recording internal structure. (Example 14).

Using Byte Offsets

The <area> elements supports specifying byte offsets in its BEGIN and END attributes, and this can be useful when using the structMap to apply structure to an archive file. Example 15, drawn from the ECHO Dep project, shows this approach to structuring an ARC file. METS, however, now supports this kind of archive file/nested file analysis in conjunction with nested file elements in the <fileSec>.

Multiple Files Manifesting a Structural Division

Although these features appear seldom to be used, METS does provide for specifying multiple files or parts of files that must be played or displayed in parallel or in sequence to manifest a structural division. Northwestern uses the <seq> element to join together two images in cases where the source item was so large that it had to be imaged in two separate segments which were then joined together to create integral derivative images. (Example 16). The California Digital Newspaper Collection uses the <seq> element to draw together multiple areas (elements) of an ALTO file that together comprise a logical division. (Example 17). NYU has used the <par> element in the case of a METS encoded multi-media presentation where an image file, a segment of an audio file, and a segment of a structured text (TEI) file must be played/displayed in parallel to manifest a logical division of the presentation. (Example 18). While METS does allow for the use of <seq> within <par> and <par> within <seq>, we have encountered no known authentic examples in the field.

Non-hierarchical Linking

The METS <structMap> only provides for hierarchical structuring. When a non hierarchical, hyperlink style structuring is desired, this may be obtained through the <structLink> element. <structLink> provide for specifying a link between two <div> elements in the same or different <structMap> elements within the same METS document. The ECHO Dep project uses this feature to model the available hyperlinks between web pages in an ARC file representing a web captures. (Example 19). The Royal Danish Library makes a rather different use. (Example 20) Where useful, it has created images of Arabic textual materials in two different orientations. It uses one structMap for the normal orientation images and one structMap for the rotated orientation images. The structLink is then used to link structural divisions in the normal orientation structMap with their rotated counterparts in the rotated orientation structMap. In addition to providing for one to one links, METS provides for many to many links conforming to the XLink extended link requirements. The British Library anticipates using this feature; however, there are no known applications of this mechanism in production. however.
element in the item list that corresponds to it. There are several tools (such as Archivist Toolkit) that are geared towards creating METS objects in conjunction with EADs.

Appendix C: Known METS Implementors
Partial List

National/State Libraries & Agencies	College/University Libraries	Public Libraries	Consortial/shared digital libraries
Library of Congress	Harvard	Boston Public Libraries (planning stages)	HathiTrust
Library & Archives Canada	Yale	Queens Borough Public Library (planning stages)	California Digital Library
British Library	MIT	L.E. Phillips Memorial Public Library, WI	California Local History Digital Resources Project
National Library of Wales	Boston College	LaCrosse Public Library	National Digital Newspaper Project
National Library of Scotland	Northeastern University	Los Angeles Public Library	Panhandle Library Access Network
German National Library	Brown University		CADLIS: Academic Digital Library of China
Staatsbibliothek, Berlin	Providence College (planning stages)		
Württembergische Landesbibliothek, Stuttgart	Southern New Hampshire University		
French National Library	Columbia University		
National Library of Austria	NYU		
Biblioteca Alexandrina, Egypt	Princeton University		
National Library of Iran	Rutgers University		
National Library of Vietnam	Florida Center for Automation (Florida state universities)		
National Library of China	University of Florida		
Chinese Ministry of Education	University of Miami	***Compiled from METS Implementation***	
National Diet Library Japan	University of Chicago	***and Profiles Registry***	
National Library of Singapore	Michigan State University	***As of October 2010***	
National Library of Australia	University of Illinois		
State Library of Victoria	Indiana University		
National Library of New Zealand	University of Texas		
Swiss National Library	Brigham Young University		
National Library of Portugal	University of California		
Ministry of Culture, Spain	University of Alberta		
National Library of Norway	Oxford University		
National Library of Finland	University of Graz		
Royal Library of Denmark	SUB Göttingen		
National Library of Sweden	Kassel University		
National Library of Latvia	Thüringer Universitäts –und Landesbibliothek, Jena		
National Library of the Netherlands	Universitätsbibibliothek Braunschweig		
National Library of Luxembourg	Australian National University		
National Library of Belgium			

Suggested Supplemental Reading [METS 2]

suggested suplemental reading for Reimagining METS
@DLF Fall Forum 2010

[Link to whitepaper]

Aligning METS with the OAI-ORE Data Model 2009-06
http://hdl.handle.net/2142/10744
McDonough, Jerome P.

Preserving Virtual Worlds Final Report; Metadata & Packaging Recommendations
section 8. Packaging Virtual Worlds pages 98-104 2010-09
http://hdl.handle.net/2142/17097
McDonough, Jerome P. et. al. -- provides model of ontology used across OAI-ORE and METS

OAI-ORE on METS wiki
https://www.socialtext.net/mim-2006/index.cgi?mets_oai_ore
https://www.socialtext.net/mim-2006/index.cgi?mcdonough
https://www.socialtext.net/mim-2006/index.cgi?habing_and_cole

Adding OAI-ORE Support to Repository Platform pages 2 - 3
http://journals.tdl.org/jodi/article/download/749/640 [PDF]
Alexey Maslov et. al.

some notes on METS as a theory of digital object description
http://tingletech.tumblr.com/post/1186443159/
Brian Tingle

XML Fever
http://dret.net/netdret/docs/wilde-cacm2008-xml-fever.html
Erik Wilde and Robert J. Glushko

Soft Issues Surrounding Industry Standard Schemas
http://xml.sys-con.com/node/40310
Sean McGrath December 28, 2001, XML Journal

HTML 5 Decentralized extensibility
http://www.w3.org/TR/html5/infrastructure.html#extensibility
http://www.w3.org/html/wg/tracker/issues/41
all elements allow data-* and _vendor-feature attributes

bad xml smells
http://inkdroid.org/journal/2010/08/25/bad-xml-smells/
Ed Summers (refactors METS/ALTO as HTML/RDF)

Miscellaneous suggestions for broadening METS web support
https://www.socialtext.net/mim-2006/index.cgi?miscellaneous_suggestions_for_broadening_mets_web_support

http://pear.ly/t7QD