Testing the Proposed METS 2.0 Data Model against Use Cases and Complementary Data Models: Presentations and Community Discussion

Bertrand Caron
Dept. of Bibliographic and Digital Information
Bibliothèque nationale de France
Quai François Mauriac
75706 Paris Cedex 13, France
Bertrand.caron@bnf.fr

Nancy J. Hoebelheinrich
Knowledge Motifs LLC
448 East Ellsworth Court
San Mateo, CA 94401, USA
+1-650-302-4493
nhoebel@kmotifs.com

Andreas Nef
Docuteam GmbH
Im Langacker 16
CH-5405 Baden-Dättwil
Switzerland
a.nef@docuteam.ch

Thomas G. Habing
Library Software Development Group
155 Grainger Engineering Library
Information Center MC-274
University of Illinois at Urbana-Champaign, USA
thabing@illinois.edu

ABSTRACT
The Metadata Encoding and Transmission Standard (METS) 1.x schema has an established community of users including academic and national libraries, archives, and museums as well as support from a number of commercial and open source tool and service vendors. While the established community of METS users has adapted systems and tools to METS expressed in XML, many in the library and archive communities are moving toward the use of newer technologies such as those of the Semantic Web and linked data for the digital content that they have been collecting. As a result, the METS Editorial Board (MEB) has been contemplating a data model for a next generation METS schema that will facilitate these kinds of technologies. The initial approach to a new METS data model aligned very closely to metadata schemes in the preservation arena, namely PREMIS, but the MEB thought it essential to test the new METS 2.0 data model against existing canonical implementations of METS, and developing complementary data models. This workshop will describe current and ongoing efforts to evaluate and further develop a new METS data model. Participants are invited to participate in the discussions, and the subsequent evaluation / refinement of a METS 2.0 data model.

General Terms
Frameworks for digital preservation; Preservation strategies and workflows; Innovative practice.

Keywords
Aggregation formats; Digital object description; Metadata Encoding and Transmission Standard, Metadata standards alignment, Digital preservation.

1. INTRODUCTION
The Metadata Encoding and Transmission Standard (METS) 1.x schema has an established community of users including academic and national libraries, archives, and museums as well as support from a number of commercial and open source tool and service vendors. While the established community of METS users has adapted systems and tools to METS expressed in XML, many in the library and archive communities are moving toward the use of newer technologies such as those of the Semantic Web and linked data for the digital content that they have been collecting. In order to accommodate the interests of this community and anticipate the needs of the established METS community, the METS Editorial Board (MEB) has been contemplating a data model for a next generation METS schema that will facilitate these kinds of technologies.

2. INITIAL APPROACH
One of the most common, and canonical uses for METS has been to facilitate the preservation of digital objects in libraries, archives and museums, often in conjunction with PREMIS. In addition, because the PREMIS standard had already been transformed into RDF, the MEB thought that an initial approach to a new METS data model should include a close alignment to other metadata schemes that were compatible and complementary, such as PREMIS and OAI-ORE. The first draft of the METS 2.0 data model, introduced in 2014, was built with these kinds of alignments in mind. Subsequent MEB and community discussion resulted in the desire to test the new METS 2.0 data model against existing canonical implementations of METS, and developing complementary data models, especially those for structuring both simple and complex digital objects.

3. APPLYING USE CASES
Use cases from two canonical implementations of METS have been developed to provide a better understanding of how the first draft of the METS 2.0 data model could be applied to both a relatively simple digital object and a more complex, 3D object. In the course of the application to the use cases, a number of issues

have arisen, such as what elements and attributes (from the XML perspective) would be important to keep or adapt from the METS 1.x schema, which could or should be used from other schemas, and the implications of the possible choices. While the two initial use cases have not yet been fully developed, the MEB would value community input upon the findings to date.

4. THE BROADER CONTEXT`

While the METS Editorial Board has been developing a next generation data model, other, similar efforts have arisen by other communities in the libraries, archives, and museum communities. The PREMIS Editorial Committee continues to work on a version 3.0 of a PREMIS ontology that promises useful symmetry with METS as does, potentially, efforts to craft the Portland Common Data Model that is being developed by the Duraspace community. As collaboration seems more productive than competition in the area of digital object aggregation and description, speakers knowledgeable about complementary data models will discuss their data models and what issues have arisen that could benefit from cross-format collaboration.

5. THE PROS / CONS OF STANDARDS ALIGNMENT

Other issues have arisen as a result of the application of use cases, and the exploration of ways to adapt the proposed METS 2.0 data model. While the choice of RDF as a means of serializing a data model permits the re-use of classes and properties from another schema, there appear to be some disadvantages to this approach that give pause. For instance, it does seem important to keep in mind the overall purpose or goal of a complementary schema to more fully anticipate the implications of the re-use of classes and properties. Given the overall purpose of a complementary schema, when and for what reasons is it more advisable to create a new class within a METS domain than re-use one from another domain? Community discussion of these issues will be solicited using specific examples that have arisen from the application of the use cases to the METS 2.0 data model.