METS in Action: Standardization and Interoperability in the Digital Library

Richard Gartner (richard.gartner@sers.ox.ac.uk)

Oxford University Library Services

Rick Beaubian (rbeaubie@library.berkeley.edu)

University of California, Berkeley

Jerome McDonough (jerome@nyu.edu)

New York University

Susan Dahl (Susan.Dahl@ualberta.ca)

University of Alberta

Brian Tingle (brian.tingle@ucop.edu)

California Digital Library

The use of SGML/XML for describing collections and digital objects has been in place since the mid-1990s (for instance, in the UK's *Internet Library of Early Journals (ILEJ)* (<http://www.bodley.ox.ac.uk/ilej/>) . By the end of that decade, the EAD ("Encoded Archival Description") had been developed to give the archivist community an encoding standard for describing archival collections digitally and for helping scholars and researchers identify and locate relevant materials in these collections. However the EAD did not address the problem of how electronic versions of the individual items comprising an archival collection could themselves be described digitally in a standard way.

Such an encoding standard needed to provide a means for inventorying the individual content files (image files, structured text files, etc.) comprising an electronic version of an archival item, applying a structure or structures to these content files, and associating relevant descriptive and administrative metadata with both structure and content. Beginning in 1997, a Digital Library Federation initiative called *Making of America II* sought to address this need for a digital object encoding standard, and out of this initiative came the MOA2.DTD; from these origins, the current METS ("Metadata Encoding and Transmission Standard") has developed.

METS is now firmly established within the digital library community, although the number of projects employing it is still relatively small. It is intended to act as a "MARC standard" for digital objects, by providing a standardized framework within which the metadata detailed above may be contained. This standardization will allow the degree of interoperability that has prevailed in the cataloguing world to become possible in the more complex environment of digital objects, and hopefully facilitate the pooling of digital resources in similar ways to the union catalogues that MARC has allowed.

This panel session will provide a short introduction to METS and its history, and show how it is currently implemented, demonstrating how it offers solutions of wide applicability to some difficulties presented by the older standards and techniques. The participants are all members of the METS Editorial Board, responsible for the maintenance of the standard. It is intended that each participant will give a short (10-15 minute presentation) and half an hour will be available for general discussion.

Rick Beaubian, Software Engineer, Digital Library Projects, University of California, Berkeley, will trace the evolution of METS from its origins in the *Making of America II* initiative to the present, and examine its progress from a relatively narrow standard primarily targetting archival and traditional library materials to one that is now also being applied to encoding audio and video content as well as to archiving websites. It will also look briefly at METS' position and application relative to other emerging content packaging standards: IMS-CP/SCORM and MPEG-21.

Richard Gartner, Pearson New Media Librarian at Oxford University Library Services, will discuss ways of integrating METS and TEI: METS is designed to allow easy integration with any XML-based documents, either by reference or by direct embedding within the METS file itself. This talk will demonstrate how TEI documents can be integrated within METS files and how the two in tandem can overcome some of the difficulties experienced when using the TEI on its own. In particular, he will cover how METS and TEI can handle images in complex objects more simply than in TEI alone, and how overlapping hierarchies can be handled neatly using multiple METS structural maps.

Jerome McDonough, Digital Library Team Leader at New York University, will discuss the application of METS to video, specifically a series of videos of performances gathered by the Hemispheric Institute. He will focus on some of the problem issues in trying to collect descriptive and structural metadata on a multi-institution project such as this, and how METS may help alleviate them.

Susan Dahl, Metadata and Cataloguing Librarian at the University of Alberta, will discuss how METS can be incorporated in a project that uses OCLC's Olive Software to digitize textual materials. Using the Peel's Prairie Provinces Project at the University of Alberta as an example (<http://peel.library.ualberta.ca/>), the presentation will

demonstrate how the XML supplied in Olive's format can be incorporated into METS documents and the benefits this offers. Also, it will feature how other descriptive, structural and administrative metadata is included, to make a complete METS document.

Brian Tingle, Content Management Designer at the California Digital Library, will discuss the Local History Digital Resources Project, which explores a model to support the creation of, and permanent public access to, standardized digital objects with associated collection guides through a single point of access. The project aims in particular to develop requirements for, and helping to select, a common digital project tool that all contributors will be required to use.

The METS home page is available at <http://www.loc.gov/standards/mets>.