

Making MODS to Linked Open Data: A Collaborative Effort for Developing MODS/RDF

Ray Denenberg

Library of Congress representative to both W3C and OASIS, Chair of the OASIS Search Web Services Technical Committee, currently involved primarily with research into the development of BIBFRAME.

Rebecca Guenther

Adjunct professor in NYU's Moving Image Archiving and Preservation Program, working remotely part-time for the Library of Congress on MODS, PREMIS, LC's Linked Data Service, and the Bibliographic Framework Transition Initiative.

Myung-Ja Han

Associate Professor of Library Administration and Metadata Librarian at the University of Illinois at Urbana-Champaign.

Jeff Mixter

Research Support Specialist at OCLC

Amy L. Nurnberger

Research Data Manager for Columbia University

Melanie Wacker

Metadata Coordinator at Columbia University Libraries and Information Services

Kathryn Pope

Head of the Scholarly Communication Program at Columbia University

Brian Luna Lucero

Digital Repository Coordinator at the Center for Digital Research and Scholarship (CDRS), Columbia University Libraries and Information Services

ABSTRACT:

Publishing library catalog records as Linked Open Data is a challenge to many libraries because there is no community-driven best practice that each individual library can easily follow and implement into its workflow.

Publishing library data as Linked Open Data is common practice for many national libraries, notably the British Library, French National Library, and the German National Library as well as for metadata aggregators and service providers, such as Europeana and the Online Computer Library Center (OCLC). However, the ways in which these institutions execute

Linked Open Data differs in many aspects. These differences are typically found in the data model used, the granularity of data, and the Linked Open Data sources used in the data, to name a few examples.

The Metadata Object Description Schema (MODS) RDF Group was formed in late 2013 as virtual working group to test and develop a MODS/RDF ontology. The group is a follow-on to an initiative of the Library of Congress. MODS was originally developed in 2002 to "give special support to cataloguing electronic resources" and as an alternative that is less detailed than, although highly compatible with, MARC21.

For this reason, is has been adopted by a wide variety of users and applications. MODS is also used as a metadata standard to which a library's traditional catalog records can be transformed while maintaining quality and granularity.

In addition, MODS has proved that its data model and rich semantics can work well in semantic Web environments :

- MODS can accommodate entity data structure introduced in FRBR

- MODS has semantics that accommodate URIs as values, in addition to strings.

The MODS/RDF Group, consisting of Librarians and programmers from a number of libraries (primarily academic institutions, OCLC and the Library of Congress), has been working to develop a MODS/RDF ontology that will allow MODS users to convert their MODS/XML metadata to RDF. The Group also hopes to publish a transformation tool, XSLT, as an end product. Since its first meeting in January 2014, the Group has created an openly viewable GitHub page (https://github.com/blunalucero/MODS-RDF) and members work together to solve the common issues in creating new, as well as using already established, Linked Data semantics that best work for the MODS data model and the information that library catalog records describe.

This presentation will share the challenges that have been encountered and the progress so far. The Group also would like to draw suggestions and recommendations for future work, especially in conjunction with other linked data work, such as BIBFRAME and Schema.org.