An Analysis of the use of MODS in Digital Repositories

Purpose
The purpose of this project is to examine the Metadata Object Description Schema (MODS) metadata scheme to determine its utility based on structure, interoperability and metadata quality.

MODS Metadata Scheme
The MODS scheme was developed by the Library of Congress to provide a schema and guidelines for encoding a resource description. MODS is a set of 20 top level elements and 56 sub-elements that can be combined with several attributes. MODS is expressed in XML, and can be used with any set of cataloging rules. The authors of MODS provide mappings and stylesheets for use with MARC, Dublin Core, RDA, and MARCXML. There are no mandatory elements, and every element except <recordInfo> can be repeated. This combination of flexibility and an ability to create highly specific records means that MODS is adaptable to a wide variety of collections and can meet the needs of a wide variety of institutions.

Methods
The MODS Implementation Registry lists 34 projects identified as using the MODS metadata scheme. Three digital repositories were selected from this Registry. Repositories were chosen based on the availability of MODS records for public view. Three repositories were examined for this project: Copac (http://copac.ac.uk/), the University of Florida Digital Collections (http://ufdc.ufl.edu/), and the Library of Congress Web Archives (http://lcweb2.loc.gov/diglib/lcwa/html/lcwa-home.html). Twenty-five records from each repository were analyzed for controlled vocabulary usage, completeness, accuracy, and consistency.

Controlled Vocabularies
The MODS scheme allows for the use of any controlled vocabulary. The “authority” attribute can be used with six of the top level elements to designate which controlled vocabulary is being used for that particular element. MODS allows for multiple usages of elements, and therefore multiple controlled vocabularies can be used for the same element. Each of the repositories made use of the Library of Congress Subject Headings (LCSH) vocabulary in at least one sample record. The Copac and University of Florida Digital Collections (UFDC) made use of the MARC Genre Term list. The Library of Congress Web Archives (LOC) also used the Thesaurus of Graphic Materials, and the LCSH name authority file. Copac had the least amount of controlled vocabulary usage amongst sample records. 8% of the Copac sample used controlled vocabulary subject terms, 72% of UFDC sample records used controlled vocabulary subject terms, and 100% of LOC sample records used controlled vocabulary subject terms. The results of this analysis show that multiple controlled vocabulary schemes can be successfully implemented in MODS records, and that the use of controlled vocabularies enhances the utility of MODS records.

Metadata Quality
Metadata quality is an essential part of building digital collections. Quality metadata enhances a user’s ability to navigate a collection, and increases interoperability among collections and software systems. The MODS element set is designed to enhance quality while allowing for flexibility. The MODS guidelines are thorough, and the amount of elements, sub elements and attributes works to limit any semantic challenges in application of elements. All three collections contained metadata of relatively good quality, and that elements were applied accurately and consistently throughout the collections. The biggest variation amongst the three collections in terms of quality is the completeness of records. The LOC collection contains the most robust records, while the UFDC and Copac have less complete records.

Conclusions
MODS is a robust metadata scheme that is designed to allow for thorough, flexible, and customizable records. This design enhances the interoperability of MODS records while still maintaining metadata quality. The examination of repositories using MODS shows that although records contain differing levels of completeness, they are highly uniform in structure, accuracy and consistency. The MODS schema may take more work to implement, but it is an excellent choice for any institution wishing to create resource descriptions.
I certify that:

· This paper/project/exam is entirely my own work.
· I have not quoted the words of any other person from a printed source or a website without indicating what has been quoted and providing an appropriate citation.
· I have not submitted this paper / project to satisfy the requirements of any other course.

Signature Carrie E. Moran
Date May 28, 2011