5-16-2006

Report of Task Force on Metadata Analysis

Margaret Mering
University of Nebraska-Lincoln, mmering1@unl.edu

Scott Childers
University of Nebraska - Lincoln, scott@emeraldfusion.net

Adonna Fleming
University of Nebraska-Lincoln, dfleming2@unl.edu

Sue Ann Gardner
University of Nebraska-Lincoln, sgardner2@unl.edu

Charity K. Martin
University of Nebraska - Lincoln, cmartin3@unl.edu

See next page for additional authors

Authors
Margaret Mering, Scott Childers, Adonna Fleming, Sue Ann Gardner, Charity K. Martin, and Judith Wolfe

Report of Task Force on Metadata Analysis

for

University of Nebraska–Lincoln Libraries

by

Margaret Mering
Scott Childers
Adonna Fleming
Sue Ann Gardner
Andy Jewell
Charity Martin
Judith Wolfe

May 16, 2006

Abstract:

This report :

1. Provides a brief overview of the different metadata schemes that are available.

2. Identifies the various metadata schemes being used to search and access information and digital content in the UNL Libraries.

3. Provides an analysis of the state of cross-searching among the various metadata schemes.

4. Provides recommendations on how UNL should decide which metadata schemes to use and when to use them.

To: Joan Giesecke, Beth McNeil & Mary Bolin
From: Margaret Mering, Scott Childers, Adonna Fleming, Sue Ann Gardner, Andy Jewell, Charity Martin, Judith Wolfe
Re: Task 2 – Metadata Analysis – Report
Date: May 16, 2006

Charge:

1. Provide a brief overview of the different metadata schemes that are available. A glossary with definitions would be helpful.

An overview of the different metadata schemes and a glossary are available in Appendices Two and Five. Appendix Three compares the elements of Dublin Core, CSDGM, and MODS. Attached are the results of work done by the Whitman Archive to identify redundant metadata in their different XML files. A listing of computer communication protocols is included in Appendix Four.

2. Identify the various metadata schemes we are using to search and access information and digital content in the UNL Libraries.

A description of the metadata schemes we are using in the Libraries is available in Appendix One.

3. Provide an analysis of the state of cross-searching among the various metadata schemes.

After researching the schemes and protocols used within the Libraries currently; the METS profile that is being created as part of the IMLS-funded Interoperability of Metadata grant; and the upload of CONTENTdm metadata into OCLC's WorldCat, our committee recognizes that universal cross-searching will be challenging to implement. We agree that a single point of access, a "Google-like" approach, would be desirable and appreciated by our patrons. As a first step, Technical Services can create doorway records (see Task 1 update) for inclusion in the catalog. These records will provide increased access to the variety of electronic materials.

To address this issue over the long term, the committee feels a new task force, with significant representation from CORS as well as from Technical Services and Digital Initiatives, would be appropriate. It is important to thoughtfully strategize solutions to both the technical and service challenges that will emerge with the creation of a central searching mechanism. Our sense is that as the complexity of the Library's electronic and print resources continues to grow, the catalog will not be the solution, but that a new solution will need to be implemented. We do not recommend developing a home-grown mechanism, but instead seek out collaborations with other institutions and/or vendors facing similar issues. An example of a current project is the eXtensible Catalog (XC). The University of Rochester has received a grant "to begin planning the requirements analysis for the development of an open-source online system to unify access to traditional and digital library resources."

A decision must be made on what path to follow before proceeding further. Below are two directions the University Libraries could choose to go if it is decided to go forward with this vision.

- For true "Google-like" full text searching, we would need to create a centralized duplicate copy of everything to be able to search text as well as the information stored in metadata fields. To do this could require a significant amount of FTE to create that database and the spidering mechanisms necessary to automate such a project. An estimate for the time necessary would be approximately a year for development, testing, and spidering. This centralized index or database with its own metadata schema would be the better solution, as queries would be answered faster and more accurately from a single database with a single schema than translating the request for full-text results from multiple databases with different schemas. Middleware may be available that would do full-text, but the results from queries might not be accurate.
- If the vision is simply to provide a single input box for searching the multiple databases using defined fields such as author, title, subject and the like, then a middleware solution is preferable. A middleware solution would be more acceptable in this case, because the limiting of searchable fields allows the translated queries into the various databases to be more structured, and the results would be more accurate. The middleware would transform the user's search into one appropriate for all the databases, no matter what their schema. This is identical to what we are doing with the "Multi-Search" option for databases we subscribe to. We can just as easily pay Innovative to create indexes for our own databases. A brief search has failed to find other institutions working on this, but an in-depth inquiry by an implementation group may be able to scratch below the surface and find partners in endeavor.

4. Provide recommendations on how UNL should decide which metadata schemes to use and when to use them.

The committee recommends that the Library creates an online storehouse of all metadata documentation decisions. Since metadata schemes are designed to address specific needs, the metadata choice of a project team will be relatively straightforward. What will require more attention is the implementation decisions of the specific schemes, and we recommend that those decisions be actively documented and posted by communities making them. For example, the CONTENTdm group is actively developing guidelines for best practices and should post these guidelines; the Center for Digital Research in the Humanities ought to provide links to its project guidelines, etc. Overall, we do not believe strict, all-purpose metadata guidelines would be helpful or practical, but instead recommend new structures to support cross-project communication.

APPENDIX ONE
METADATA SCHEMES USED IN THE UNL LIBRARIES

CSDGM (Content Standard for Digital Geospatial Metadata)

- Currently, Fleming and Mering are working in partnership with the Conservation and Survey Division of the School of Natural Resources to create and update metadata for CSD's geospatial datasets
- In the future, the Libraries will support the creation of CSDGM metadata for geospatial datasets developed by other areas on campus as part of the GIS program.

Dublin Core/VRA

- All collections in CONTENTdm use Dublin Core or VRA.
- Examples include:
    - College of Architecture, Decorative Arts
    - Department of Art and Art History, Art History Survey 101
    - Eloise Kruger Collection
    - Historic Textiles
    - Larsen Tractor Test and Power Museum
    - Omaha Indian Heritage Project (under construction)
    - Platte River Basin.

EAD

- Walt Whitman Archive
- Archives and Special Collections finding aids.

MARC

- The Libraries catalog is MARC-based and resides on an Innovative Interfaces (III) system.

METS

- METS is likely to be more widely applied in the future. Currently it is a part of an interoperability of metadata grant that the Walt Whitman Archive is participating in.

TEI

- All projects in the Center for Digital Research in the Humanities and the Electronic Text Center that have a textual component use the TEI standard for encoding
- Examples include:
    - Willa Cather Archive
    - Journals of the Lewis and Clark Exhibition
    - Walt Whitman Archive
    - Good Person: Excerpts from the Yoruba Proverb Treasury
    - Birds of Nebraska
    - Plains Humanities Alliance projects.

APPENDIX TWO
SELECTED LIST OF METADATA SCHEMES

Content Standard for Digital Geospatial Metadata (CSDGM)
aka FGDC Metadata Standard
http://www.fgdc.gov/

The Content Standard for Digital Geospatial Metadata (CSDGM), version 2 (FGDC-STD-001-1998) is the US Federal Metadata standard. The Federal Geographic Data Committee originally adopted the CSDGM in 1994 and revised it in 1998. According to Executive Order 12096 all Federal agencies are ordered to use this standard to document geospatial data created as of January, 1995. The standard is often referred to as the FGDC Metadata Standard and has been implemented beyond the federal level with State and local governments adopting the metadata standard as well.

Currently, the US is revising the CSDGM to be compliant with ISO 19115. Each nation can craft its own profile of ISO 19115 with the requirement that it include the 13-core element. The FGDC is currently leading the development of a US Profile of the international metadata standard, ISO 19115.

DDI—Data Documentation Initiative
http://www.icpsr.umich.edu/DDI/

Somewhat similar to FGDC, DDI is a metadata scheme for datasets in the social sciences, often collected in the form of surveys that have columns of data, with corresponding codebooks that explain what the data mean. The idea is to make these data more sharable, interoperable, and machine analyzable. Originally written in SGML, DDI is now in XML. Similar to the home pages of other schemes, at the DDI Web site, there is information on the scheme, its history, its governance and development, and project using DDI.

Dublin Core
http://dublincore.org
http://dublincore.org/documents/dcmi-terms/

Dublin is designed to be used with all-disciplines, in contrast to some other schemes, which are designed for particular disciplines or types of materials. Dublin Core grew out of an OCLC/National Center for Supercomputing Applications (NCSA) Metadata Workshop held in March 1995 in Dublin, Ohio. The scheme's name is based on the location of that workshop. Its elements represent the CORE or minimal elements necessary to describe resources. The Dublin Core consists of fifteen elements. They are Title, Creator, Subject, description, publisher, contributor, date, type, format, identifier, source, language, relation, coverage, and rights. Audience is recommended as the sixteenth element. No elements are mandatory and all elements are repeatable.

Dublin Core can be either simple or qualified. Qualifiers further explain elements. The element Date has a number of qualifiers. For example dateAccepted, dateCopyrighted, and

dateSubmitted. The Western Trails and Western Waters projects used dateOriginal and dateDigital.

EAD—Encoded Archival Description
http://lcweb.loc.gov/ead/

Encoded Archival Description (EAD) is a standard for encoding archival finding aides using XML. It is maintained in the Network Development and MARC Standards Office of the Library of Congress in partnership with the Society of American Archivists. Development of the EAD DTD began with a project initiated by the University of California, Berkeley, Library in 1993.

III MetaSource
http://www.iii.com/pdf/lit/eng_metasource.pdf

From the III Web site: "Innovative's MetaSource is a suite of tools that allows libraries to effectively manage their digital collections. This includes digital object storage, crawling external collections, and full support for metadata schemes such as Dublin Core."

Also from the III Web site, about MetaSource: "Libraries need a multi-faceted solution for describing and digitizing media collections. These collections are often in different formats (*e.g.*, some electronic and some paper), challenging libraries with the formidable task of handling and describing their disparate collections. Libraries need options to store, crawl, index, and describe these collections, as well as the choice to either integrate them into the traditional bibliographic catalog or maintain separate collections. Emerging standards also need to be accommodated so that these collections are interoperable.

"MetaSource is made up of three components: Millennium Media Management, XML Harvester, and MetaData Builder. Millennium Media Management creates and stores media objects such as images, sound files, and audio files. It also includes a Copyright and Access component to handle the complex licensing and copyright issues of digital collections. The XML Harvester gathers XML records from any server; it then parses and creates records on the Innovative system. MetaData Builder stores XML in the metadata scheme of choice. Together, these tools create a comprehensive digital library management strategy.

Millennium Media Management XML Harvester MetaData Builder"

LOM—Learning Object Metadata
http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

LOM is often created for online courses that are delivered via various courseware packages such as Blackboard, WebCT, or Desire2Learn. The objective in using the LOM scheme is to facilitate re-use of course materials by more than one instructor, no matter which courseware is used.

Among others, IEEE, the Department of Defense and an organization called Advanced Distributed Learning (ADL) are developing the LOM scheme. ADL's scheme is called Sharable

Content Object Reference Model (SCORM) (http://www.adlnet.org). IEEE's scheme is referred to as IEEE LOM.

IMS is an organization involved in developing specifications for LOM that software and metadata developers can follow if they choose to do so. At the IMS Web site, there is also some documentation on what elements are necessary for learning object metadata, and some examples of what LOM, encoded in XML, looks like. IEEE has also drafted standards for LOM (http://ltsc.ieee.org).

MADS—Metadata Authority Description Schema
http://www.loc.gov/standards/mads/

MADS is intended to complement MODS as the authorities component. From the MADS Web site, "The Library of Congress …, with interested experts, has developed the Metadata Authority Description Schema (MADS), an XML schema for an authority element set that may be used to provide metadata about agents (people, organizations), events, and terms (topics, geographics, genres, etc.). MADS was created to serve as a companion to the Metadata Object Description Schema (MODS). As such, MADS has a relationship to the MARC 21 Authority format, as MODS has to MARC 21 Bibliographic—both carry selected data from MARC 21.

MARC
http://www.loc.gov/marc/marc.html

MARC is the acronym for MAchine-Readable Cataloging. It was developed by the Library of Congress in 1965 for the interchange of bibliographic information. The MARC record consists of three main components: Leader, Directory, and Variable Fields. The Leader and Directory contain information about the record itself. The Variable Fields contain metadata about the resource being described. The Network Development and MARC Standards Office at the Library of Congress and the Standards and the Support Office at the Library and Archives Canada maintain the MARC format. The ALCTS/LITA/RUSA Machine-Readable Bibliographic Information Committee (MARBI) is the committee within the American Library Association responsible for developing official ALA positions on the MARC format's development.

METS version 1.5—Metadata Encoding and Transmission Standard
http://www.loc.gov/standards/mets/

Developed at the Library of Congress, METS, in part, is a preservation scheme for metadata to enable continuing access to digital collections that will inevitably require occasional migrations to newer computing platforms. METS is a standard for encoding descriptive, administrative and structural metadata regarding digital objects, expressed using the XML schema language. METS is appropriate for complex digital objects, such as illustrated multivolume sets of primarily textual items. METS builds on the XML object model used in the Making of America II project, and creates a standard for an XML document that is a package, containing descriptive and technical metadata for a digital object, as well as links to the object itself.

At the METS Web site, it states: "The metadata necessary for successful management and use of digital objects is both more extensive than and different from the metadata used for managing

collections of printed works and other physical materials. While a library may record descriptive metadata regarding a book in its collection, the book will not dissolve into a series of unconnected pages if the library fails to record structural metadata regarding the book's organization, nor will scholars be unable to evaluate the book's worth if the library fails to note that the book was produced using a Ryobi offset press. The same cannot be said for a digital version of the same book. Without structural metadata, the page image or text files comprising the digital work are of little use, and without technical metadata regarding the digitization process, scholars may be unsure of how accurate a reflection of the original the digital version provides."

A METS document contains seven main sections: 1) METS header, 2) descriptive metadata, 3) administrative metadata, 4) file section, 5) structural map, 6) structural links, 7) behavior (this is optional and can include executable functions).

For examples of projects using METS, see the METS Implementation Registry at the METS Web site (http://www.loc.gov/standards/mets/). The California Digital Library uses METS for its eScholarship editions (http://texts.cdlib.org/escholarship/).

MIX (NISO Metadata for Images in XML Schema)
http://www.loc.gov/standards/mix/

The Library of Congress' Network Development and MARC Standards Office, in partnership with the NISO Technical Metadata for Digital Still Images Standards Committee and other interested experts, is developing an XML schema for a set of technical data elements required to manage digital image collections. The schema provides a format for interchange and/or storage of the data specified in the NISO Draft Standard Data Dictionary: Technical Metadata for Digital Still Images (Version 1.2). This schema is currently in draft status and is being referred to as "NISO Metadata for Images in XML (NISO MIX)". MIX is expressed using the XML schema language of the World Wide Web Consortium. MIX is maintained for NISO by the Network Development and MARC Standards Office of the Library of Congress with input from users.

MODS version 3.1—Metadata Object Description Schema
http://www.loc.gov/standards/mods/

Also created at the Library of Congress, MODS is an extension of METS, and is intended to complement other metadata formats. While METS can carry structural, administrative and descriptive metadata, MODS is primarily for MARC (i.e. descriptive metadata). From the MODS Web site: "As an XML schema, … MODS is intended to be able to carry selected data from existing MARC 21 records as well as to enable the creation of original resource description records. It includes a subset of MARC fields and uses language-based tags rather than numeric ones, in some cases regrouping elements from the MARC 21 bibliographic format." It includes 19 top level elements such as titleInfo, genre, language, subject and physicalDescription, all of which have assigned attributes, and some of which also contain subelements.

From the MODS Web site, "MODS could potentially be used as follows:

- as a Z39.50 Next Generation specified format (http://www.loc.gov/z3950/agency/zing/zing.html)
- as an extension schema to METS (Metadata Encoding and Transmission Standard)
- to represent metadata for harvesting
- for original resource description in XML syntax
- for representing a simplified MARC record in XML
- for metadata in XML that may be packaged with an electronic resource."

ONIX
http://www.editeur.org

ONIX is an e-commerce standard for the book and serial industries. The standard was developed and is maintained by EDItEUR, the Book Industry Communication (UK), and the Book Industry Study Group (US) and other user groups. ONIX includes data elements similar to MARC, such as ISBN, author, title, edition, publisher, publishing dates and subjects. It also provides access to table of contents, cover images, publisher descriptions and reviews.

TEI—Text Encoding Initiative
http://www.tei-c.org/

Text Encoding Initiative (TEI) is a standard for encoding documents. Although its use is meant to be interdisciplinary, it tends to be associated with the humanities. The standard was developed in 1987. It is maintained by a consortium of institutions and projects worldwide.

The heart of the TEI specification is a Document Type Definition (DTD) that prescribes the elements to be used in marking up a TEI encoded electronic text. The current version of TEI guidelines is referred to as P4. P5 is under development.

VRA Core version 3.0
http://www.vraweb.org/vracore3.htm

The VRA Core was developed by the Visual Resources Association for use by slide librarians and curators of visual materials collections who often have both works of visual art in their collections and images that document them. In VRA Core, separate records are created for the original object and for the surrogate. The Visual Information Access (VIA) Project at Harvard uses VRA principles. "Eye of the Beholder" by Robin Wendler, in *Metadata in Practice*, describes Harvard's VIA in detail.

APPENDIX THREE
METADATA COMPARISON

Dublin Core Elements	CSDGM Elements	MODS Elements
Title	Title	Title Info
Creator	Originator	Name
Subject	theme keyword; place keywords	Genre, subject
Description	Abstract	Table of contents; abstract
Publisher	Publisher	Name
Contributor	Contributor	Name
Date	Time period of content Time period of information	Physical description
Type	Type of source Media (paper map)	Type of Resources genre
Format	Geospatial Data Presentation Form (ArcInfo files)	Type of Resource
Identifier		Identifier
Source	Source Citation	Origin Info
Language		Language
Relation		Related Item
Coverage	Spatial Reference Information	Target Audience
Rights	Use constraints	Access Condition

APPENDIX FOUR
COMPUTER COMMUNICATION PROTOCOLS

In our research we identified a handful of computer communication protocols that are worth mentioning. These protocols could possibly be used in creating interoperable solutions for searching or transferring metadata from one schema to another.

OAI (Open Archives Initiative)
http://www.openarchives.org/

The Open Archives Initiative Protocol for Metadata Harvesting defines a mechanism for harvesting XML-formatted metadata from repositories. The protocol does not provide a mechanism for harvesting data (content) that is not encoded in XML. The protocol also does not mandate the means of association between that metadata and related content. Since many clients may want to access the content associated with harvested metadata, data providers may deem it appropriate to define a link in the metadata to the content. The mandatory Dublin Core format provides the identifier element that can be used for this purpose. (From http://www.openarchives.org/documents/FAQ.html.)

OpenURL
http://www.niso.org/committees/committee_ax.html

The OpenURL standard is a syntax to create web-transportable packages of metadata and/or identifiers about an information object. Such packages are at the core of context-sensitive or open link technology.

The OpenURL is needed because conventional web links do not take into account the identity of the user: they take all users to the same target. This causes some problems. For example, when more than one institution provides access to copies of the same electronic article, the link from citation to full text should resolve to a copy that is accessible to the user. Since different users have access to different digital libraries, the link should resolve in a user-specific fashion. In order to do this, a link must be able to:

1. Package metadata and identifiers describing the information object.
2. Send this package to a link-resolution server or resolver.

If this resolver is aware of the user's context, it is able to take into account the identity of the user when resolving the metadata into specific targets. (From http://www.niso.org/committees/committee_ax.html.)

Z39.50
http://www.niso.org/z39.50/z3950.html

A computer protocol that can be implemented on any platform, defines a standard way for two computers to communicate for the purpose of information retrieval. A Z39.50 implementation

enables one interface to access multiple systems providing the end-user with nearly transparent access to other systems. (From http://www.niso.org/z39.50/z3950.html.)

APPENDIX FIVE
GLOSSARY

TERM	DEFINITION
Best practices	Practices beyond the scope of application rules that illustrate the best ways to implement a given infrastructure component.
Catalog	Listing of resources for a collection that providing a record of individual items and collections for easier access.
CONTENTdm	Digital collection management software.
Crosswalk	A mapping of the elements, semantics, and syntax from one metadata scheme to another.
Data set (or Dataset)	A collection of data variables that have been derived from a single data source. The data set contains, or can contain, multiple types of content, where each type is associated with an object. A named collection of logically related data items arranged in a prescribed manner.
Discovery	Enable a person to find an object for which an element (i.e. author, title, subject) is known, show what the library has, and assist in the choice of object.
Element	An element is a property of a resource. As intended here, "properties" are attributes of resources—characteristics of a resource, such as a title, publisher, or subject. Elements are formally defined terms which are used to describe attributes and properties of a resource. An "element" is also the term applied to tag names in XML-based metadata schemes.
Harvesting	A technique for automatically extracting metadata from individual repositories and collecting it in a central catalog to facilitate search interoperability.
Interoperability	The ability of multiple systems, using different hardware and software platforms, data structures, and interfaces, to exchange and share data.
Middleware	Software that connects two or more different programs or databases, passing and translating data requests from the requesting application to the application storing the data and then doing the same for the output.
Migration	The movement of one metadata record or more from one location (e.g. IRIS) to another (e.g. CONTENTdm).
Navigation	The user's ability to discover collections or web content and the underlying hyperlinks that create a seamless discovery process.

PREMIS	PREMIS (Preservation Metadata: Implementation Strategies) is an international working group founded by RLG and OCLC to define implementable, core preservation metadata, with guidelines and recommendations for management and use. In May 2005, PREMIS released *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group*, which can be found here: http://www.oclc.org/research/projects/pmwg/premis-final.pdf.
Registry of Digital Masters	The Digital Library Foundation/OCLC Registry of Digital Masters provides a central place for library staff to search for and find digitally preserved materials. As such, the Registry broadens access to your organization's publicly available digital books and journals. The purpose for a digital registry: avoid duplication of effort, optimize available funding, improve access to digital material, create standards–metadata, digitization, access , develop best practice in the field, more access at less cost–collaboratively build a greater mass of digital materials than we could achieve individually.
XML	Extensible Mark-up Language is an encoding syntax that assists in the creation, retrieval, and storage of documents. It consists of a tag structure that identifies specific information within a document. Unlike HTML, XML is not limited to a specific set of tags, because a single tag set would not adapt to all documents or applications that may use XML.
XML schema	XML schemas express shared vocabularies and allow machines to carry out rules made by people. They provide a means for defining the structure, content, and semantics of XML documents in more detail.
Z39.50	Z39.50 is a communication standard that overcomes problems of multiple database searching. The protocol facilitates search and retrieval of information from databases.

APPENDIX FIVE
SUGGESTED FURTHER READING

We reviewed the following when compiling the report:

Calhoun, Karen. (2006) "The changing nature of the catalog and its integration with other discovery tools, final report, March 17, 2006, prepared for the Library of Congress." 52 p. http://www.loc.gov/catdir/calhoun-report-final.pdf

"Distributed interoperable metadata registry," published in the December 2001 issue of *D-Lib Magazine*: http://www.dlib.org/dlib/december01/blanchi/12blanchi.html

Kurth, Martin, David Ruddy, and Nathan Rupp. (2004) "Repurposing MARC metadata: using digital project experience to develop a metadata management design." *Library Hi Tech* 22(2): 153-165.
http://dspace.library.cornell.edu/bitstream/1813/1457/1/Kurth_MARC_ALCTS2005_Public.ppt
Also available at: http://www.library.cornell.edu/tsweb/metadata/p153.pdf

Larsgaard, Mary Lynette. (2005) "Metacataloging of digital geospatial data." *The Cartographic Journal* 42(3): 231-237.

Mann, Thomas. (2006) "The changing nature of the catalog and its integration with other discovery tools, final report, March 17, 2006, prepared for the Library of Congress: a critical review." 26 p. http://www.guild2910.org/AFSCMECalhounReviewREV.pdf

Marcum, Deanna B. (2005) "The future of cataloging." Presented at the EBSCO Leadership Seminar, January 16, 2005, Boston, Mass.
http://www.loc.gov/library/reports/CatalogingSpeech.pdf

"Mellon Grant funds planning analysis for future online services." (2006) *University of Rochester News*, April 14, 2006. http://www.rochester.edu/news/show.php?id=2518

Westbrooks, Elaine L. (2005) "Remarks on metadata management." *OCLC Systems & Services* 21(1): 5-7.

Additional pertinent sources include:

List of metadata crosswalks
http://www.loc.gov/marc/marcdocz.html

University of Virginia's digital library glossary
http://www.lib.virginia.edu/digital/reports/dl_terminology_uva.htm

Fedora
http://www.lib.virginia.edu/digital/resndev/fedora.html

Final Report for the AMeGA (Automatic Metadata Generation Applications) Project
http://www.loc.gov/catdir/bibcontrol/lc_amega_final_report.pdf

Access level for serials cataloging
http://www.loc.gov/catdir/access/accessrecord.html

Minimal level catalog record standards
http://www.loc.gov/marc/bibliographic/nlr/

Consortium for the Computer Interchange of Museum Information
http://www.cimi.org/