8-2016

Connecting Access from Metadata to MARC

Amy Leigh Allen
University of Arkansas, Fayetteville, ala005@uark.edu

Mary A. Gilbertson
University of Arkansas, Fayetteville, mag@uark.edu

Deborah E. Kulczak
University of Arkansas, Fayetteville, dkulczak@uark.edu

Recommended Citation

Connecting Access from Metadata to MARC

Amy Leigh Allen
Mary A. Gilbertson
Deborah E. Kulczak
University of Arkansas Libraries

Abstract

Making a connection between digital repositories and library catalogs is an issue many libraries and archives face, especially when dealing with items that traditionally received full cataloging records. This poster will explore different workflows for reusing metadata to create catalog records, connecting the two resources together. In order to have one interface to search both the older physical items and newer digital items, the archives and technical services department of the University of Arkansas Libraries worked together to create different workflows to save time and eliminate the need for double entry, using the open source software MarcEdit and XML Notepad. We created two processes for distinct materials in different repositories. The first process transforms metadata for undergraduate theses in DSpace from Dublin Core to MARC. The second process transforms metadata from music concert recordings in Islandora from MODS to MARC.

MODS	11
Concert Recordings	225
Islandora	2012

Dublin Core	21
Honors Theses	642
DSpace	2011

Background – Concert Recordings

The Music Department at the U of A has been recording concerts performed by their students and faculty since the 1960s and has consistently given a copy of these recordings to the University Libraries since the early 1980s. In late 2012, the Music Department determined that they would no longer produce CDs of these recordings but instead would give the Libraries digital files, which they wanted streamed on the web. The Archives, in conjunction with campus IT services, had developed a digital repository using an Islandora platform which could fulfill this need. Islandora offered simple Dublin Core or MODS metadata, so the librarians and archivists working on the project opted to use MODS. However, brief MARC records had already been created in the Libraries' catalog for nearly 2,600 concert recordings dating back to 1963, and library faculty wanted to continue cataloging concerts so that all of them could be searched in the same interface.

Concert Recordings Workflow

University Archivist Amy Allen, Music Cataloging Unit Head Deborah Kulczak, and then Head of the Performing Arts and Media Library (PAM) Lora Lennertz approached the problem with the goal of creating a single XML record for each recording that could serve as the basis for both the Islandora metadata and the catalog record. In the resulting workflow, the original MODS XML record is created by staff in PAM with the open source program XML Notepad. They can then use the Islandora batch import feature to upload the metadata, along with the digital files, into that repository.

Creating the catalog records is a two-step process. The XML files are first batch converted from MODS XML to MARC with MarcEdit software. Its MARC Tools XML function comes with an existing MODS => MARC converter (XSLT stylesheet) that we could use with just a single modification. Then the MARC file is further customized and loaded into our Innovative catalog by means of a locally-created load table. (For institutions lacking this option, XSLT could be used for the entire process.)

Background – Honors Theses

Undergraduate Honors Theses were maintained historically by the individual colleges.

In 2010, the University Archivist began working with the Engineering Librarian, who had previously collected theses in digital format. The University Archives worked with Library and campus IT services to develop a DSpace platform and then a Vireo platform – an ETD management system developed by the Texas Digital Library. In 2011, student began to upload their theses directly into Vireo. Librarians and archivists wanted honors theses available through the Libraries catalog similar to the graduate theses.

Metadata

In both cases, the University Archivist and Technical Services faculty members worked together to define required metadata fields with specific definitions for each project. The honors theses use the metadata schema developed by the Networked Digital Library of Theses and Dissertations (NDLTD), which is based on Dublin Core metadata while the Concert Recordings use MODS metadata.

MODS to MARC Crosswalk

MODS Field	MARC (datafield tag)
<typeOfResource>	LDR
<location><shelfLocator>	099
<titleInfo><subTitle>	245
<titleInfo type="alternative"><title>	246
<originInfo><dateIssued keyDate="yes">	260 $c, 518
<note>	500
<tableOfContents>	505
<subect><topic>	690
<name type="personal"><namePart>	700 $a
<name type="personal">	
<role><roleTerm type ="text">	700 $e
<name type="corporate"><namePart>	710
<location><url>	856

Concert Recordings in Islandora:
http://hdl.handle.net/10975/islandora:181420

Undergraduate Theses in Dspace:
https://uarkive.uark.edu/xmlui/handle/10826/51

Libraries Catalog: http://library.uark.edu/

DSpace to MARC Crosswalk

DSpace Field	MARC (datafield tag)
dc.title	245
dc.creator	100
dc.date.created	500
dc.contributor.advisor	690
dc.contributor.committeemember	690
thesis.degree.level	690
thesis.degree.name	690
thesis.degree.department	690
thesis.degree.discipline	690
thesis.degree.grantor	690

Honors Theses First Procedure

Developed with University Archivist Amy Allen and Cataloger Elaine Dong.
- Student uploaded thesis into Vireo, creating metadata
- Thesis published from Vireo to DSpace
- Archivist exported batch of metadata from DSpace to a .csv file
- Cataloger edited .csv file in Excel
- Cataloger used the Mail Merge function in Word to turn Excel file into a .txt file, edited fields according to AACR2 and saved with UTF-8 encoding.
- Cataloger used MarcEdit to convert the .txt file to a MARC file
- Cataloger uploaded MARC file to catalog.

Honors Theses Second Procedure

The Second procedure still originated from DSpace, but then used MarcEdit's OAI Harvester with modified XSLT script instead of Excel and Mail Merge to convert the data. This method was tested by Amy Allen and Cataloger Mary Gilbertson with the help of a student programmer to edit the XSLT crosswalk that converts the Dublin Core based metadata to the designated MARC coding. Staff could then export the metadata directly from Dspace using MarcEdit and import the resulting file in to the catalog.

WARNING CHALLENGES AHEAD

Before the customized XSLT was finished, the campus decided to adopt Bepress for the campus wide institutional repository. The theses have been migrated from Dspace to Bepress and the Libraries' faculty are working on a new workflow for cataloging the theses imported into Bepress.

Under Construction
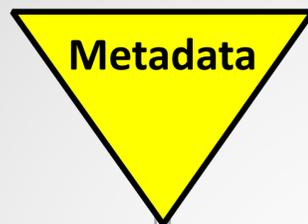
Suggested Resources

- Brian E. Surratt and Dustin Hill, "ETD2MARC: A semiautomated workflow for cataloging electronic theses and dissertations," *Library Collections, Acquisitions & Technical Services* 28 (2004): 205-223.
- Sevim McCutcheon, Michael Efeyche, Margaret Beecher Maurer, and Joshua Nickerson, "Morphing metadata: Maximizing Access to Electronic Theses and Dissertations," *Library Hi Tech*, 26, no. 1 (2008): 41-57.
- Michael Boock and Sue Kunda, "Electronic Thesis and Dissertation Metadata Workflow at Oregon State University Libraries," *Cataloging & Classification Quarterly* 47, no. 3/ 4 (2009): 297-308.

You Have Arrived at the Catalog

Conclusion

There are many options and different tools for reusing metadata to create catalog records. In both of these workflows, MarcEdit was found to be a very useful tool that offered options for harvesting and converting both Dublin Core and MODS metadata. MarcEdit is an open source tool, which is easily installed without programming knowledge. In order to customize the crosswalks, a knowledge of XML and XSLT will be needed. It may also be necessary to have a familiarity with load tables in order to load the records into the catalog. With a few simple tools, you can travel the path from metadata to MARC.