



# Metadata re-write

Moving Avalon metadata  
from XML to RDF



# Avalon Media System



Browse

Search

Browse By

Format

Moving Image 11

Main contributor

Date

Collection

Unit

Language

You searched for: Format > Moving Image

Start Over

« Previous | 1 - 10 of 11 | Next »

Sort by Relevance

10 per page

1. Avalon Sprint #100 Demo



Date: 2016-06-03

Main contributors: The Avalon Media System Team

2. Avalon Sprint #102 Demo



Date: 2016-06-24

Main contributors: The Avalon Media System Team

Summary: The Avalon team's demo for sprint 102 includes: showing our new waffle.io board, walking through our current design for a playlist marker and demonstrating RuboCop in our development branch

3. Avalon Sprint #103 Demo



Date: 2016-07-09

Main contributors: The Avalon Media System Team

Summary: 103 focused on: • Packaging bug release 5.0.1 • Adding new play list features for 5.1, of note is our markers feature • Fedora 4 research for Avalon 6.0 release 103's demo is our new

Browse

Search

**Editing playlist**

[View Playlist](#) [Delete Playlist](#)

**Playlist Details**

Name: Greatest Moments of Pro Wrestling  
Description: The suplexes! The below the belt punches!  
Visibility: Private

**Playlist Items**

1	Macho Man Randy Savage Highlight Reel
2	The Undertaker Goes Down
3	The Hulkster Rides
4	Beyond the Ring

[Avalon Media System Project Website - Contact Us](#)

Avalon Media System Release 4.0.1

## Avalon Sprint #100 Demo

**Date**

2016-06-03

**Main contributor**

The Avalon Media System Team

**Collection**

Avalon Sprint Demos

**Unit**

University Archives

Share

Browse Manage Content Manage Groups Manage Selected Items (0) Playlists

Search

Manage file(s)

Associated bitstreams

Resource description

Metadata about the item

Structure

Organization of resources

Access Control

Who can access the item

Preview

How end users see the item

Delete this item

Avalon Sprint Demos > Avalon Sprint #100 Demo

## Resource description

Fields followed by an asterisk (\*) are required.

### Bibliographic ID ⓘ

Catalog Key ▾

Import

### Title\* ⓘ

Avalon Sprint #100 Demo

### Publication date\* ⓘ

2016-06-03

### Main contributor(s) ⓘ

The Avalon Media System Team

+

### Creation date ⓘ

### Summary ⓘ

# Avalon Metadata

## Before Release 6.0

- Fedora 3
- XML files attached to object
  - Description (MODS)
  - Fedora relationships (RELS-EXT)
  - Technical
  - Rights
  - Workflow
  - Sections
  - Structure

## Release 6.0

- Fedora 4
- RDF properties describing object
  - Fedora relationships
  - Technical
  - Rights
  - Sections
- XML files attached to object
  - Description (MODS)
  - Workflow
  - Structure

# Metadata Migration

XML -> RDF

Encoding -> Connecting

Within .xml file attached to Fedora object:

```
<mods:name type="personal">  
  <mods:namePart>Grainger, Porter.</mods:namePart>  
  <mods:role>  
    <mods:roleTerm authority="marcrelator" type="text">creator</mods:roleTerm>  
  </mods:role>  
</mods:name>
```

->

As property on Fedora object:

```
<Fedora object> <dcterms:creator>  
<http://id.loc.gov/authorities/names/n89670927> .
```

# Metadata Migration

XML -> RDF

Encoding -> Connecting

RDF

**<subject>** **<predicate>** **<object>** .

<Fedora object>

<dcterms:creator>

<<http://id.loc.gov/authorities/names/n89670927>> .

---



# Avalon MODS

Title	titleInfo / title
Date	originInfo / dateIssued
Main contributor	name@usage="primary" / namePart && role = Creator
Creation date	originInfo / dateCreated
Summary	abstract
Contributor	name/namePart && role = Contributor
Publisher	originInfo / publisher
Language	language / languageTerm
Physical description	relatedItem@type="original" / physicalDescription / extent
Related item	relatedItem@displayLabel / location / url
Genre	genre
Subject	subject / topic
Time period	subject / temporal
Location	subject / geographic
Permalink	location / url@access="object in context"
Terms of use	accessCondition@type="use and reproduction"
Other identifier	relatedItem@type="original" / identifier
Note	note     tableOfContents

And RDF Conversion.

- Survey #1 (Group Focus and Title Element): <http://goo.gl/forms/GzHtKl2iXk> (Results: MODS Survey #1 Results)
- MODS Title Individual Institution Usage And RDF Conversion

**MODS elements as RDF Collaboration Documents (Detailed page: [Collaboration Documents](#))**

1. MODS title (two tabs)
2. MODS name (two tabs)
3. MODS typeOfResource
4. MODS genre
5. MODS originInfo (three tabs)
6. MODS language
7. MODS physicalDescription
8. MODS abstract
9. MODS tableOfContents
10. MODS targetAudience (only BPL and UNC-CH mapped this)
11. MODS note (two supported options likely)
12. MODS subject
13. MODS classification (only BPL and Columbia mapped this)
14. MODS identifier
15. MODS physicalLocation
16. MODS accessCondition
17. MODS recordInfo
18. MODS series and collections

**Initial MODS XML to RDF resources:**

- UC Santa Barbara mapping information [MODS and RDF: UCSB Application Profile](#)
- UC San Diego mapping information: [MODS and RDF: UCSD Application Profile](#)
- MODSRDFV2: <https://github.com/blunalucero/MODS-RDF>
- Amherst College initial start of a mapping: <https://gitlab.amherst.edu/acoburn/acdc-metadata-profile>
- Emory University: [starter RDF mapping](#) (multiple namespaces), [use cases](#), and [open questions](#). Also, [working list of RDF Namespaces of interest](#)
- MODS RDF Primer: <http://www.loc.gov/standards/mods/modsrdf-primer.html>

## MODS and RDF Descriptive Metadata Subgroup

The screenshot displays a Google Sheet titled "MODS OriginInfo as RDF Collab" comparing MODS XML and MODS RDF representations. The sheet is viewed in a browser window, with a Duraspace interface visible in the background on the left.

**MODS element**

	A	B
12	<b>MODS XML Representation:</b>	<b>MODS RDF Representation:</b>
13	<pre>&lt;!-- Origin place --&gt; &lt;mods:originInfo&gt;   &lt;mods:place&gt;     &lt;mods:placeTerm type="text"&gt;Boston, Mass.&lt;/mods:placeTerm&gt;   &lt;/mods:place&gt; &lt;/mods:originInfo&gt;</pre>	<pre>&lt;self&gt; loc:pup "Boston, Mass."</pre>
14	<pre>&lt;!-- Origin place, publisher, edition --&gt; &lt;mods:originInfo eventType="publication"&gt;   &lt;mods:place&gt;     &lt;mods:placeTerm type="text"&gt;Boston, Mass.&lt;/mods:placeTerm&gt;   &lt;/mods:place&gt;   &lt;mods:publisher&gt;published by John P. Soule&lt;/mods:publisher&gt;   &lt;mods:edition&gt;3rd edition&lt;/mods:edition&gt; &lt;/mods:originInfo&gt;</pre>	<pre>&lt;self&gt; loc:pup "Boston, Mass." &lt;self&gt; dcterms:publisher "published by John P. Soule" &lt;self&gt; bibframe:edition "3rd edition"</pre>
15	<pre>&lt;!-- Multiple events --&gt; &lt;mods:originInfo eventType="manufacture"&gt;   &lt;mods:publisher&gt;Steven Anderson&lt;/mods:publisher&gt;   &lt;mods:place&gt;     &lt;mods:placeTerm type="text"&gt;[Atlanta, GA]&lt;/mods:placeTerm&gt;   &lt;/mods:place&gt; &lt;/mods:originInfo&gt;</pre>	<pre>&lt;self&gt; loc:pup "[Atlanta, GA]" &lt;self&gt; dcterms:publisher "Steven Anderson" &lt;self&gt; loc:pup "Boston, Mass." &lt;self&gt; dcterms:publisher "John P. Soule"</pre>
16		
17	<pre>&lt;mods:originInfo eventType="publication"&gt;   &lt;mods:place&gt;     &lt;mods:placeTerm type="text"&gt;Boston, Mass.&lt;/mods:placeTerm&gt;   &lt;/mods:place&gt;   &lt;mods:publisher&gt;John P. Soule&lt;/mods:publisher&gt;   &lt;mods:edition&gt;3rd edition&lt;/mods:edition&gt;</pre>	<pre>&lt;self&gt; loc:pup "[Atlanta, GA]" &lt;self&gt; dcterms:publisher "Steven Anderson" &lt;self&gt; loc:pup "Boston, Mass." &lt;self&gt; dcterms:publisher "John P. Soule"</pre>

**Initial MODS XI**

- UC Santa
- UC San D
- MODSRD
- Amherst C
- Emory Un
- Namespac
- MODS RE

## MODS and RDF Descriptive Metadata Subgroup

My Drive - Google Drive | My Drive - Google Drive | Avalon MODS to RDF - Goo... | Metadata re-write - Google... | +

https://docs.google.com/spreadsheets/d/1TY5G\_9Jl6Ql-Aq6c3Cs54BH5E5sRu30mhnz6GMub0U/edit#gid=0

Avalon MODS to RDF ☆

File Edit View Insert Format Data Tools Add-ons Help Last edit was made on February 24 by jennifer young

Comments Share

ihardes@iu.edu

123 Arial 10 B I A

	A	B	C	D	E
1	bf2: http://id.loc.gov/ontologies/bibframe/ bibframe: http://bibframe.org/vocab/ dbpedia: http://dbpedia.org/ontology/ dc: http://purl.org/dc/elements/1.1/ dcterms: http://purl.org/dcterms/ edm: http://www.europeana.eu/schemas/edm/ loc: http://id.loc.gov/vocabulary/identifiers/ owl: http://www.w3.org/2002/07/owl# relator: http://id.loc.gov/vocabulary/relators/				
2	<b>Form field label</b>	<b>MODS element</b>	<b>RDF property</b>	<b>Notes</b>	
9	Contributor	name/namePart (role/roleTerm set to "Contributor")	dc:contributor	requires URI for value - migration would need to resolve text names to URIs or create local URIs for names; helpful to have form with Questioning Authority-style autocomplete to work off of LCNAF. Changing from MODS/RDF mapping to dc:creator because too many names will not map - easier path forward to bring this in as text value.	
10	Genre	genre	edm:hasType		
11	Publisher	originInfo/publisher	dcterms:publisher		
12	Date Created	originInfo/dateCreated@encoding="edtf"	dcterms:created		
13	Date Issued	originInfo/dateIssued@encoding="edtf"	dcterms:issued		
14	Abstract	abstract	dcterms:abstract		
15				MODS and RDF Subgroup mapping these to ISO639-2 but we're using MARC Language codes, so migration would end up with URI values like:	

Generic model | IU examples | NU examples

Explore

# Avalon MODS to RDF

# Migration and Identifiers

Not only migrating descriptive metadata but set of identifiers that connect this object and its description to external sources and past forms.

- Fedora 3 identifiers (avalon:pid)
- MODS identifiers (OCLC, LCCN, music publisher, local)
- Permalink (PURL)
- Bibliographic Import ID (Sirsi Catalog Key)

# Migration and Identifiers

- Fedora 3 identifiers (avalon:pid)
  - <Fedora Object> <<http://www.w3.org/ns/prov#wasDerivedFrom>>  
<[http://\[fedora3URL.com\]/avalon:19920](http://[fedora3URL.com]/avalon:19920)> .
- MODS identifiers (i.e., OCLC, local)
  - <Fedora Object> <<http://dbpedia.org/ontology/oclc>> "ocm05309912" .
  - <Fedora Object> <<http://id.loc.gov/vocabulary/identifiers/local>> "09-010.08-F" .
- Permalink (PURL)
  - <Fedora Object> <<http://purl.org/dc/elements/1.1/identifier>>  
<<http://purl.dlib.indiana.edu/iudl/media/415p290961>> .
- Bibliographic Import ID (Sirsi Catalog Key)
  - <Fedora Object> <<http://www.w3.org/2002/07/owl#sameAs>>  
<<http://dlib.indiana.edu:9000/iucateextract?operation=searchRetrieve&version=1.1&query=cql.serverChoice%3d%27^C2231159%27&maximumRecords=1>> .

But wait, there's more to consider!



# Hyrax

- New development focus for Hydra community
- Based on Sufia/IR Hydra Head
- Already in Fedora 4 and using RDF for description



Avalon Hyrax property comparison

File Edit View Insert Format Data Tools Add-ons Help Last edit was made on February 22 by anonymous

Comments Share

Form field label	MODS element	Hyrax RDF property	MODS To RDF MAPPING
Bibliographic ID	relatedItem@type="original"/identifier	dcterms:bibliographicCitation? NO	loc:local
Bibliographic ID Label	relatedItem@type="original"/identifier@type		
Other Identifier	relatedItem@type="original"/identifier		
Other Identifier Label	relatedItem@type="original"/identifier@type		local = loc:local; lccn = loc:lccn; issue number = loc:issue-number; matrix number = loc:matrix-number; music publisher = loc:music-publisher; video recording identifier = loc:videorecording-identifier; oclc = dbpedia:oclc
Title	titleInfo/title	dc:title	dcterms:title
Creator	name@usage="primary"/namePart (role/roleTerm set to "Creator")	dc:creator	relator:cre
Contributor	name/namePart (role/roleTerm set to "Contributor")	dc:contributor	relator:ctb
Genre	genre		edm:hasType
Publisher	originInfo/publisher	dc:publisher	dcterms:publisher
Date Created	originInfo/dateCreated@encoding="edtf"		dcterms:created
Date Issued	originInfo/dateIssued@encoding="edtf"	dcterms:created	dcterms:issued
Abstract	abstract	dc:description	dcterms:abstract

+ Descriptive Technical Non-matching Hyrax properties Explore

## Avalon MODS vs Hyrax vs MODS and RDF Subgroup

48%

Hyrax mapping only matches 12 out of 25 fields used in Avalon.

Push development to implement MODS and RDF Subgroup's mapping in Hyrax?

# Questioning Authority

## Autocomplete Saving String and URI

Authority: OCLC FAST

Subauthority: PersonName

Oclc organization **required**

Cornell University

+ Add another Oclc organization

Oclc person **required**

ezra

- Benson, Ezra Taft
- Cornell, Ezra, 1807-1874**
- Ezra (Biblical figure)
- Gannett, Ezra S. (Ezra Stiles), 1801-1871
- Ibn Ezra, Abraham ben Meir, 1089-1164
- Ibn Ezra, Moses, approximately 1060-approximately 1139
- Meeker, Ezra, 1830-1928
- Pound, Ezra, 1885-1972
- Sims, Ezra, 1928-2015
- Stiles, Ezra, 1727-1795

Requirements

- Describe your work

Visibility

- Open Access** Everyone. Check out SHERPA/RoMEO for specific publishers' copyright policies if you plan to patent and/or publish your Demo Work in a journal.
- Institution** Restrict access to only users and/or groups from Institution
- Embargo**
- Lease**
- Private** Only users and/or groups that have been given specific access in the "Share With" section.

I have read and agree to the **Deposit Agreement**

Cancel Save

Lynette Rayle,  
Linked Data in Hyrax,  
LDCX 2017

# Questioning Authority

## Getting more from other authorities?

**Oclc person** [Cornell, Ezra, 1807-1874](#)

(source: [http://dbpedia.org/resource/Ezra\\_Cornell](http://dbpedia.org/resource/Ezra_Cornell) )

Birth: 1807-01-11

Death: 1874-12-09

Ezra Cornell (January 11, 1807 – December 9, 1874) was an American businessman, politician, philanthropist and educational administrator. He was the founder of Western Union and a co-founder of Cornell University. He also served as President of the New York Agriculture Society and as a New York state Senator.

**Oclc person** [uri](http://id.worldcat.org/fast/409667)

Lynette Rayle,  
[Linked Data in Hyrax](#),  
LDCX 2017

14425 records

Extensions: undefined ▾ RDF ▾

Show as: rows records Show: 5 10 25 50 records

« first < previous 1 - 10 next > last »

	Genre 2	Genre 2 URI	Genre 3	Genre 3 URI	Genre 4	Genre 4 URI	Topical Subject	Topical Subject Headings 1 URI	Topical S
9	Identification photographs <small>Choose new match</small>	<a href="http://id.loc.gov/vocabulary/graphicMaterials/tgm000665">http://id.loc.gov/vocabulary/graphicMaterials/tgm000665</a> <small>edit</small>					Automobiles <small>Choose new match</small>	<a href="http://id.loc.gov/vocabulary/graphicMaterials/tgm000665">http://id.loc.gov/vocabulary/graphicMaterials/tgm000665</a>	People <small>Choose new mat</small>
4	Identification photographs <small>Choose new match</small>	<a href="http://id.loc.gov/vocabulary/graphicMaterials/tgm005292">http://id.loc.gov/vocabulary/graphicMaterials/tgm005292</a>					Mountains <small>Choose new match</small>	<a href="http://id.loc.gov/vocabulary/graphicMaterials/tgm006835">http://id.loc.gov/vocabulary/graphicMaterials/tgm006835</a>	Buildings <small>Choose new mat</small>

# Side Note!

[Data Wrangling with OpenRefine](#)

Scholars' Commons Workshop  
Wells Library W144

Tuesday, April 18, 2-3:30pm

```
143 <oi:record>
144   <oi:header>
145     <oi:identifier>cushman4</oi:identifier>
146     <oi:datestamp>2017-04-06</oi:datestamp>
147   </oi:header>
148   <oi:metadata>
149     <qdc:qualifieddc xmlns:qdc="http://www.bl.uk/namespaces/oi_dcq/"
150       xmlns:dc="http://purl.org/dc/elements/1.1/"
151       xmlns:dcterms="http://purl.org/dc/terms/"
152       xsi:schemaLocation="http://www.bl.uk/namespaces/oi_dcq/
153         http://www.bl.uk/schemas/qualifieddc/oi_dcq.xsd">
154       <dc:identifier>http://purl.dlib.indiana.edu/iudl/archives/cushman/P01415</dc:identifier>
155       <dc:source>http://purl.dlib.indiana.edu/iudl/archives/cushman/thumbnail/P01415.jpg</dc:source>
156       <dcterms:created>1938-09-03</dcterms:created>
157       <dc:description>A-11= Burbank Garden Scenes</dc:description>
158       <dc:title>A-11=in Burbank Garden Santa Rosa , Calif.</dc:title>
159       <dcterms:spatial>United States, California, Santa Rosa, 38.440429, -122.7140548, http://maps.google.com/maps?q=38.440429,-122.7140548</dcterms:spatial>
160       <dc:subject>Luther Burbank Home and Gardens</dc:subject>
161       <dc:subject>Gardens, Flowers, Dwellings, Shrubs, Trees</dc:subject>
162       <dc:format>Landscape photographs</dc:format>
163       <dc:rights>http://rightsstatements.org/vocab/InC-EDU/1.0/</dc:rights>
164       <dcterms:provenance>Indiana University, Bloomington</dcterms:provenance>
165       <dc:type>Image</dc:type>
166       <dc:language>English</dc:language>
167       <dcterms:accessRights>IN COPYRIGHT - EDUCATIONAL USE PERMITTED</dcterms:accessRights>
168       <dc:creator>Cushman, Charles Weever, 1896-1972(photographer)</dc:creator>
169       <dcterms:isPartOf>Charles W. Cushman Photograph Collection</dcterms:isPartOf>
170     </qdc:qualifieddc>
171   </oi:metadata>
172 </oi:record>
```

DPLA feed for Cushman

# IU Migration Plan - Descriptive Metadata

- Similar to Avalon, using MODS to RDF Subgroup recommendations
  - Also probably incorporating within Hyrax, either using changes from Avalon work or customizing to accommodate MODS in different collections
- Linked Data reconciliation per collection
  - Obtain URIs where possible for fields like subjects, locations, and names using authorized controlled vocabularies
  - Depending on collection, mapping for DPLA might better inform content migrated into Fedora 4



Questions? Comments!

jlhardes@iu.edu

@jlhardes #dlbb