Library of Congress – Subject Heading Import – Notes
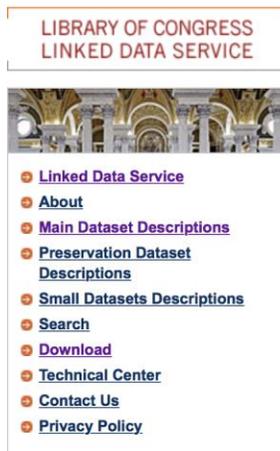
This document gives a quick overview of importing the Library of Congress Subject Heading file into Axaem.

Installation

This application uses the 0DX application to perform the XML import. If you have not configured the 0DX application for your database, do so now. When creating the files, make sure the RAWXML and RAWXMLAT files are set to FMS type 9, as they will probably go over the default 4Gig file size limit.

Getting Started

First download the subject file from the Library of Congress web site at http://id.loc.gov/download/. Download the 'LC Subject Headings (MADS/RDF only)' file:

Download

Bulk Downloads

Bulk downloads, serialized as RDF/XML, Turtle, or N-triples, are available for the authorities and vocabularies. These may be downloaded as compressed files (ZIP format), with the exception of the LC Name Authority File, which is GZipped. Unless otherwise noted, each download contains MADS/RDF and SKOS/RDF representations of the data.

Read more »

Vocabulary/Authority	Date	Size	Download
LC Children's Subject Headings	27 Oct 2014	425 (kb)	nt
LC Children's Subject Headings	27 Oct 2014	198 (kb)	rdf
LC Children's Subject Headings	27 Oct 2014	155 (kb)	ttl
LC Demographic Group Terms	4 Sep 2015	20 (kb)	nt
LC Demographic Group Terms	4 Sep 2015	53 (kb)	rdfxml
LC Demographic Group Terms	4 Sep 2015	22 (kb)	ttl
LC Genre/Form Terms	27 Oct 2014	621 (kb)	nt
LC Genre/Form Terms	27 Oct 2014	363 (kb)	rdf
LC Genre/Form Terms	27 Oct 2014	298 (kb)	ttl
LC Name Authority File (MADS/RDF only)	7 Apr 2016	2645374 (kb)	nt
LC Name Authority File (SKOS/RDF only)	25 Apr 2016	921387 (kb)	nt
LC Name Authority File (MADS/RDF only)	5 Apr 2016	2837099 (kb)	rdfxml
LC Name Authority File (SKOS/RDF only)	25 Apr 2016	990906 (kb)	rdfxml
LC Medium of Performance Thesaurus for Music	27 Oct 2014	515 (kb)	nt
LC Medium of Performance Thesaurus for Music	27 Oct 2014	286 (kb)	rdf
LC Medium of Performance Thesaurus for Music	27 Oct 2014	250 (kb)	ttl
LC Subject Headings (MADS/RDF only)	19 May 2016	165174 (kb)	nt
LC Subject Headings (SKOS/RDF only)	19 May 2016	37526 (kb)	nt
LC Subject Headings (MADS/RDF only)	19 May 2016	178282 (kb)	rdfxml
LC Subject Headings (SKOS/RDF only)	19 May 2016	40929 (kb)	rdfxml

LIBRARY OF CONGRESS LINKED DATA SERVICE

- Linked Data Service
- About
- Main Dataset Descriptions
- Preservation Dataset Descriptions
- Small Datasets Descriptions
- Search
- Download
- Technical Center
- Contact Us
- Privacy Policy

Uncompress the file and put it somewhere on the server where it will be accessible to APPX. Make a note of the location and file name.

Next, you have to enter that name and location into the Axaem Parameter file. This setting is on the third page of the Parameter file in the 'Path to XML Import File':

Finally, you have to designate one of your authorities as the Library of Congress. On the 'Process' tab, run 'Authority File Maintenance', and designate one of the authorities as 'Library of Congress'.

You can only designate one at a time, so if you have the wrong one designated, you must first remove it from the incorrect one before changing the new one.

Running the Import

The first step in importing the data is to run 'Load Library of Congress Subject Headings'. This step reads the XML file and 'flattens' the data into some work files (SUBJLC, SUBJLCEL and SUBJLCND). This step will run for several hours with no user intervention required, and no progress indicator. Be patient. When it is complete, a report of the number of records created and unhandled XML attributes is printed. If any errors were encountered (such as data truncation), they will be reported.

Sample output:

```
NC Archives - Mars Conversion                                    Page:   1
             !Import Subject/Node                                04/05/17 15:08
                                                                          SPF

          Date Full                    Parent Rid  Attribute
================================================================================
Import Started at: 2017-04-05 03:08:32.41pm
End Import:        2017-04-05 10:11:06.44pm RC:
Value was truncated: #text
NODE            (A closet drama is a play that is not intended to be performed onstage, but read by a solitary reader or, somet

Value was truncated: #text
NODE            p. 163 (Juvenile Delinquency Film (Delinquent Movie, Hoodlum Drama, J. D. Film). A substantial number of youth
```

```
NC Archives - Mars Conversion                                    Page:   3
             !Import Subject/Node                                04/05/17 23:32
                                                                          SPF

          Date Full                    Parent Rid  Attribute
================================================================================
Unhandled XML Attributes Encountered:
rdf:resource
xml:lang
```

The 'rdf:resource' and 'xml.lang' attributes are normally ignored. Anything else should be investigated and any errors should be addressed before continuing.

If all is well, continue the import process by running the 'Process Loaded Data' option. This will use the SUBJLC* files from the previous step and create a hierarchy of files suitable for importing into Axaem (LCSUBJ* files).

A progress counter is displayed while it runs. There are typically 400,000+ records to process so this will also take a while to complete, but not as long as the first step.

The system will have to translate the Library of Congress Subject Types to Axaem Subject Types.

After this step is complete, you will get a report of the Subject Types that are not defined in Axaem:

```
================================================================================
The following Subject Types were found in the Library of Congress data, but
do not have corresponding types in AXAEM. This must be corrected before
the data can be imported.

Authority
```

Use this report to add or change the necessary Access Points in Axaem:

Access Point Type	Ead Tag	MARC Code	Loc Type
AE, Corporate Name	corpname	710	Authority
AE, Personal Name	persname	700	NameTitle
AE, Uncontrol. Title	subject	740	ComplexSubject
Corporate Name	corpname	610	CorporateName
Deprecated	cdata	010	DeprecatedAuth…
Faceted Topical Term	subject	654	Temporal
Family Name	famname	600	FamilyName
Function	function	657	ConferenceName
Genre/Form	genreform	655	GenreForm
Geographic N			Geographic
Local Subject			
Local-Corpor.			
Meeting Name			
Occupation			
Personal Nam			PersonalName
Topical Term			Topic
Uncontrolled	subject	653	Variant
Uniform Title	title	630	Title

Access Point Types dialog:
Type: *Geographic Name
EAD Tag: *geogname
MARC Code: *651
LoC Type: Geographic
OK   Cancel

It may not be necessary to set up all new Types, you may be able to simply enter the Library of Congress type into the 'Loc Type' field of existing types.

Failure to set up all the types will cause the transfer step to fail when it encounters a type it cannot convert to Axeam. Also note that Utah has its own MARC codes. These must be in sync with the Axaem MARC codes.

Once you have defined the necessary Types, you can begin transferring subjects into Axaem.

Transfer Library of Congress headings to Axaem

This step will allow you to transfer the LoC subject headings into Axaem:

Library of Congress Subjects Lookup List					

Subjects Filter Options

	From	To	
Keyword: robots	Date Added: MM/DD/CCYY — MM/DD/CCYY —	Max Num of Rec: 100	
Subject Title:	Date Chg: MM/DD/CCYY — MM/DD/CCYY —	Sort Field:	
LoC Subject Id	Date Dep: MM/DD/CCYY — MM/DD/CCYY —	Sort Order:	Search
In Axaem: [?]			

Select the Records to add   Records Found: 63

Subj RID	Subj Code	Already in AX...	Subject Search Title	Date Added	Date Changed	Date De
848627	sh97006962	☑	Autonomous robots	09/09/1997		
606523	sh2015002051	☐	Autonomous robots--Law and legislation	08/18/2015		
854393	sh98004829	☐	Evolutionary robotics	08/03/1998		
592674	sh2011002448	☐	Human–robot interaction	05/16/2011	08/31/2011	
491384	sh2007005901	☑	Military robots	07/31/2007	05/01/2008	
592369	sh2011001982	☐	Military robots--Mechanical scalability	04/20/2011	07/19/2011	
777842	sh88005238	☑	Mobile robots	08/25/1988		
573943	sh2010102191	☐	Mobile robots--Congresses	02/02/2010		
437900	sh00000121	☑	Parallel robots	04/05/2000	05/22/2000	
856478	sh98007352	☐	Personal robotics	11/12/1998		
722419	sh85114625	☐	Robot camera	02/11/1986		
805320	sh92004682	☑	Robot hands	08/24/1992	07/24/2015	
722420	sh85114626	☐	Robot industry	02/11/1986	08/07/1997	
582353	sh2010111343	☐	Robot industry--United States	07/13/2010		
459875	sh2003001169	☐	Robot Sumo (Game)	10/02/2003	10/02/2003	
722421	sh85114627	☐	Robot vision	02/11/1986	04/22/1988	
582354	sh2010111344	☐	Robot vision--Congresses	07/13/2010		
776990	sh88004266	☐	Robot wrists	07/13/1988	09/30/1992	
568887	sh2010011967	☐	Robotic exoskeletons	06/30/2010	08/03/2010	
462758	sh2003005867	☑	Robotic pets	03/25/2003	05/08/2003	
595276	sh2011005839	☑	Robotic sailboats	02/02/2012		
722422	sh85114628	☑	Robotics	02/11/1986	02/07/1990	
761842	sh86004454	☐	Robotics consultants	07/21/1986		
825870	sh95001354	☐	Robotics in medicine	02/17/1995		
860515	sh99003865	☐	Robotics in sports	05/13/1999		
769015	sh87004057	☐	Robotics laboratories	06/25/1987		

Review Subject in AXAEM	Add/Update AXAEM	Select	Cancel

The usual search options apply.

To transfer a subject heading to Axaem, just double click it, or single click and then click 'Add/Update AXAEM'. You can also select a range of rows, then click 'Add/Update AXAEM' to import several subject headings at once.

If the LoC subject has modifiers, those will be automatically imported as well. They will be flagged as 'Reference Only' and will not have their own modifiers (or other child data) imported. If you select a reference only subject and use the 'Add/Update AXAEM' then its modifiers (and other child data) will be imported and it will no longer be considered a reference only subject.

Once a heading is imported into Axaem, you can click the 'Review Subject in Axaem' button to review its data in the normal Axaem input.

Generally speaking, subject headings imported into Axaem cannot be changed in Axaem. You can add additional data (modifiers, MARC codes, etc), but the specific data that was imported from LoC cannot be changed or deleted (although the entire subject can be deleted).

Refresh Subjects from Library of Congress

Use this to batch update the data in Axaem with the latest subject data from Library of Congress. Only subjects that were originally imported will be affected, and only if the date changed on the Library of Congress data is more recent that the LoC date changed in Axaem. Reference only subjects are not automatically updated.

This step will first confirm that the necessary Subject Types are set up in Axaem. If any are missing, a report is printed and the job cancels. Add the missing types (see 'Running the Import' above) and try again.