

Finding Aids: MARC to EAD Conversion Process

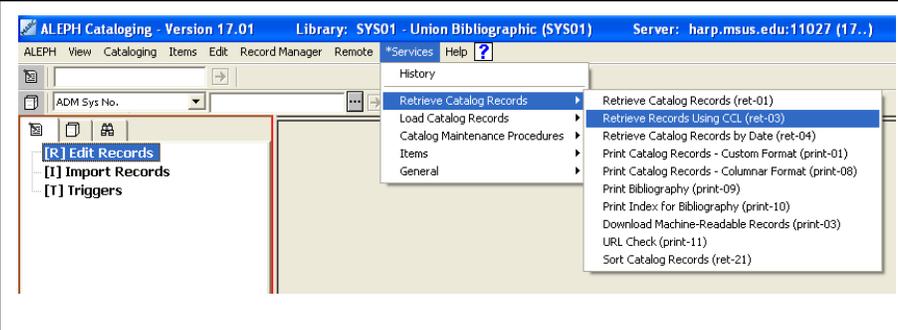
The steps outlined in this document show how to convert a MARCXML file exported from Aleph into an EAD file.

Take note of the Aleph system number for the record. You will need that number to complete this process.

Two requests, available from the *Services* menu in the Aleph cataloging module, must be sent to the MnPALS server. The first request retrieves the bibliographic record. The second request downloads that record into a MARC XML file.

After downloading the MARC XML file from Aleph, the MARCXML file is copied and pasted into an Oxygen XML document, and a transformation scenario changes the file into EAD.

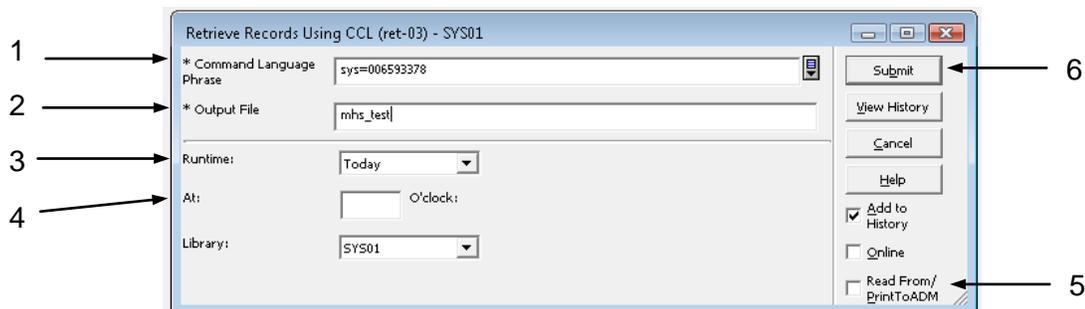
1. Retrieve Record Using CCL (ret-03):

<p>Select the <i>Services</i> menu.</p> <p>Select <i>Retrieve Catalog Records</i>.</p> <p>Select <i>Retrieve Records Using CCL (ret-03)</i>.</p>	 <p>The screenshot shows the ALEPH Cataloging interface. The title bar reads 'ALEPH Cataloging - Version 17.01 Library: SYS01 - Union Bibliographic (SYS01) Server: harp.msus.edu:11027 (17..)'. The menu bar includes 'ALEPH View Cataloging Items Edit Record Manager Remote *Services Help ?'. The 'Services' menu is open, showing options like 'Retrieve Catalog Records', 'Load Catalog Records', 'Catalog Maintenance Procedures', 'Items', and 'General'. The 'Retrieve Catalog Records' option is expanded, showing a sub-menu with 'Retrieve Catalog Records (ret-01)', 'Retrieve Records Using CCL (ret-03)', 'Retrieve Catalog Records by Date (ret-04)', 'Print Catalog Records - Custom Format (print-01)', 'Print Catalog Records - Columnar Format (print-08)', 'Print Bibliography (print-09)', 'Print Index for Bibliography (print-10)', 'Download Machine-Readable Records (print-03)', 'URL Check (print-11)', and 'Sort Catalog Records (ret-21)'. The 'Retrieve Records Using CCL (ret-03)' option is highlighted.</p>
--	--

Aleph will open a dialog box. Use the common command language¹ “SYS=[Aleph System Number]” to request the specific MARC record you wish.

Complete the retrieval service in the following manner:

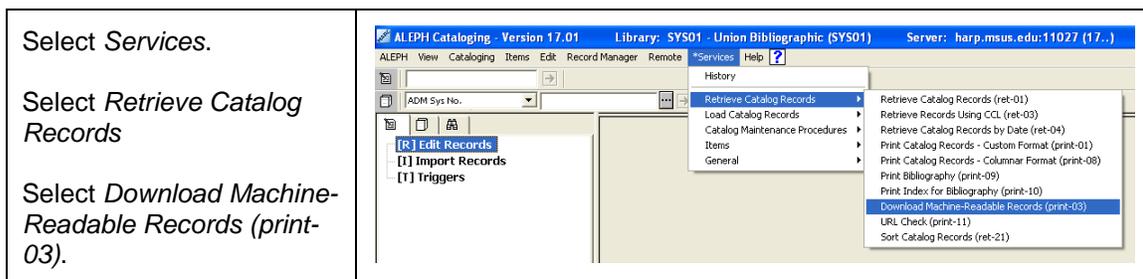
¹ The key to additional common command language queries is available at: http://tables.prod20.mnpals.net/u20_1/sys01/tab/tab11_word.



1. **Command Language Phrase:** Enter *SYS=Aleph System Number*. Use the specific Aleph System Number for the bibliographic record you want to download.
2. **Output File:** Type in a unique filename that is easy for you to remember. Enter the filename in lower case. Do not use any spaces, punctuation marks, or special characters. Underscores and hyphens may be included.
3. **Run Time:** Select *Today*. Other choices include *Tomorrow* or *DayAfterTomorrow*.
4. **At:** Enter the time when you want your request to be processed. You can express a specific hour from 00 (midnight) to 23 (11 p.m.). Leave the time blank and the server will immediately process your request.
5. **Read From/Print to ADM:** Keep the box unchecked.
6. Click **Submit** to send your request to the server.

2. Download Machine-Readable Records (print-03)

This second service uses the output file created in the first service to download the MARC record into a MARCXML file that will be used for the EAD conversion.



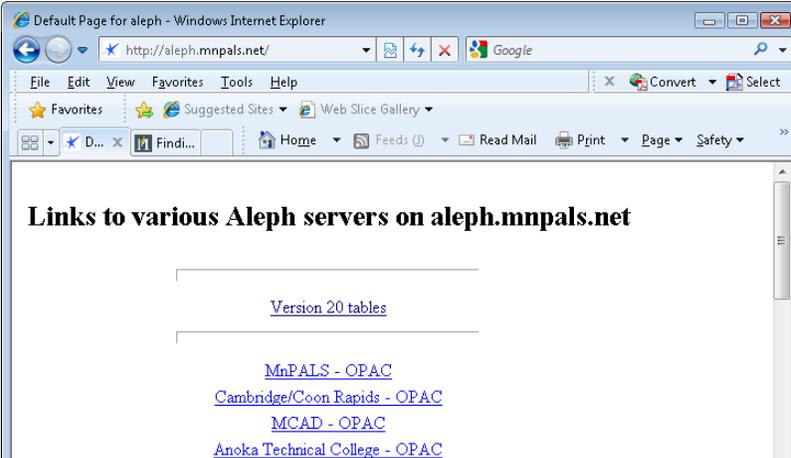
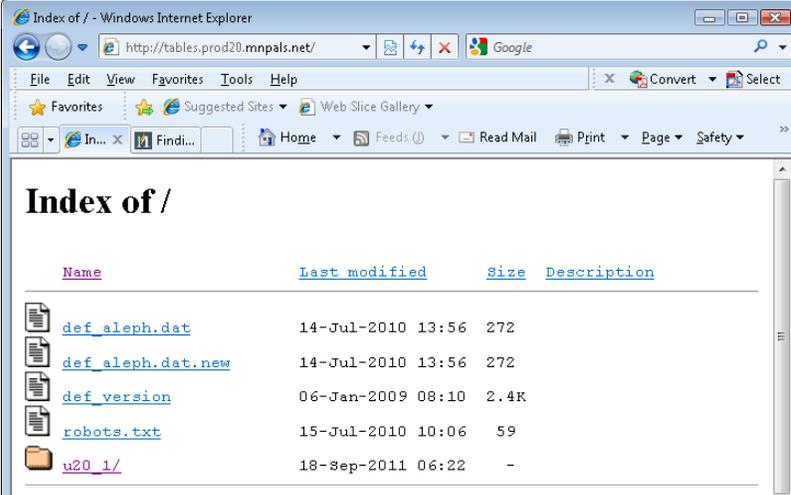
Aleph will open a dialog box that asks you to specify the fields and file format of the bibliographic record you wish to download. Complete the box in the following manner:

1. **Input File:** Enter the same name of the file that you specified as the output file in the first service request.
2. **Output File:** Enter another unique file name that is easy to remember (use no spaces and no special characters other than the underscore or hyphen).
3. **Field 1:** Enter *ALL*.
4. **Format:** Select *MARC XML* for the file format.
5. **Run Time:** Select *Today*, *Tomorrow*, or *DayAfterTomorrow*.
5. **At:** Leave blank.
6. **Read From/Print to ADM:** Make certain the box unchecked.
7. **Submit.**

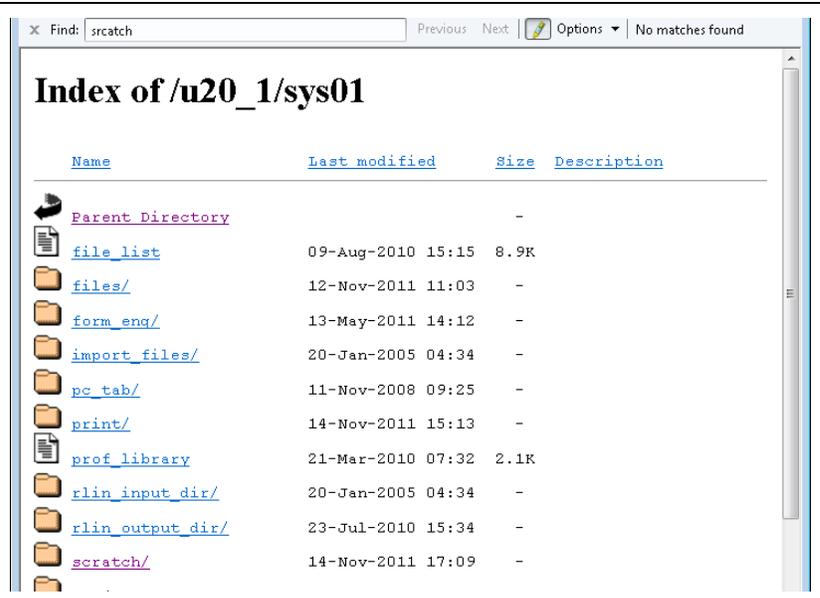
3. Open the Downloaded File

Open your browser to http://tables.prod20.mnpals.net/u20_1/sys01/scratch/ and find the file you downloaded.

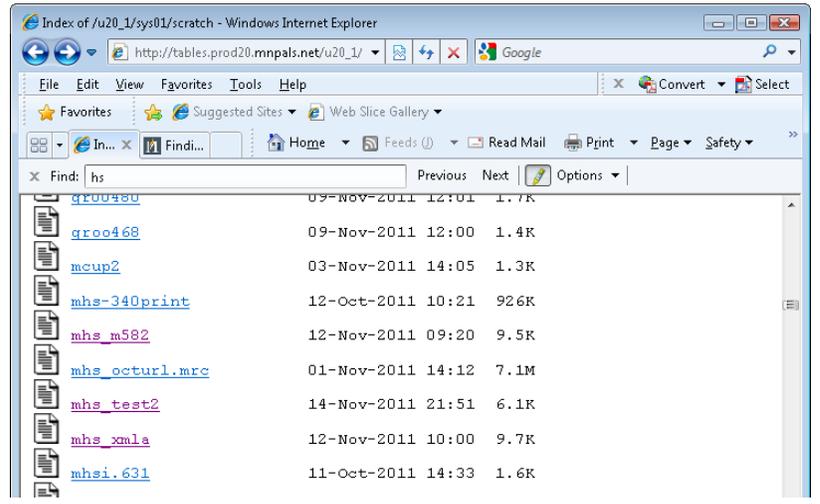
The scratch directory contains the output files created by the download service. You can also navigate through the MnPALS server by following these steps:

<p>Open your browser to http://aleph.mnpals.net/.</p> <p>Select <i>Version 20 tables</i>.</p>	 <p>Default Page for aleph - Windows Internet Explorer</p> <p>http://aleph.mnpals.net/</p> <p>Links to various Aleph servers on aleph.mnpals.net</p> <p>Version 20 tables</p> <p>MnPALS - OPAC</p> <p>Cambridge/Coon Rapids - OPAC</p> <p>MCAD - OPAC</p> <p>Anoka Technical College - OPAC</p>																								
<p>Open the <i>u20_1</i> directory.</p>	 <p>Index of / - Windows Internet Explorer</p> <p>http://tables.prod20.mnpals.net/</p> <p>Index of /</p> <table border="1"> <thead> <tr> <th>Name</th> <th>Last modified</th> <th>Size</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td>def_aleph.dat</td> <td>14-Jul-2010 13:56</td> <td>272</td> <td></td> </tr> <tr> <td>def_aleph.dat.new</td> <td>14-Jul-2010 13:56</td> <td>272</td> <td></td> </tr> <tr> <td>def_version</td> <td>06-Jan-2009 08:10</td> <td>2.4K</td> <td></td> </tr> <tr> <td>robots.txt</td> <td>15-Jul-2010 10:06</td> <td>59</td> <td></td> </tr> <tr> <td>u20_1/</td> <td>18-Sep-2011 06:22</td> <td>-</td> <td></td> </tr> </tbody> </table>	Name	Last modified	Size	Description	def_aleph.dat	14-Jul-2010 13:56	272		def_aleph.dat.new	14-Jul-2010 13:56	272		def_version	06-Jan-2009 08:10	2.4K		robots.txt	15-Jul-2010 10:06	59		u20_1/	18-Sep-2011 06:22	-	
Name	Last modified	Size	Description																						
def_aleph.dat	14-Jul-2010 13:56	272																							
def_aleph.dat.new	14-Jul-2010 13:56	272																							
def_version	06-Jan-2009 08:10	2.4K																							
robots.txt	15-Jul-2010 10:06	59																							
u20_1/	18-Sep-2011 06:22	-																							
<p>Scroll down the list and open the <i>sys01</i> directory.</p>	 <table border="1"> <tbody> <tr> <td>ssu50/</td> <td>23-Jul-2010 13:48</td> <td>-</td> <td></td> </tr> <tr> <td>sys01/</td> <td>23-Jul-2010 13:47</td> <td>-</td> <td></td> </tr> <tr> <td>sys10/</td> <td>01-May-2011 21:33</td> <td>-</td> <td></td> </tr> <tr> <td>sys19/</td> <td>23-Jul-2010 15:43</td> <td>-</td> <td></td> </tr> <tr> <td>sys30/</td> <td>23-Jul-2010 13:47</td> <td>-</td> <td></td> </tr> <tr> <td>sys40/</td> <td>11-Nov-2008 09:30</td> <td>-</td> <td></td> </tr> </tbody> </table>	ssu50/	23-Jul-2010 13:48	-		sys01/	23-Jul-2010 13:47	-		sys10/	01-May-2011 21:33	-		sys19/	23-Jul-2010 15:43	-		sys30/	23-Jul-2010 13:47	-		sys40/	11-Nov-2008 09:30	-	
ssu50/	23-Jul-2010 13:48	-																							
sys01/	23-Jul-2010 13:47	-																							
sys10/	01-May-2011 21:33	-																							
sys19/	23-Jul-2010 15:43	-																							
sys30/	23-Jul-2010 13:47	-																							
sys40/	11-Nov-2008 09:30	-																							

Scroll down and open the scratch directory.



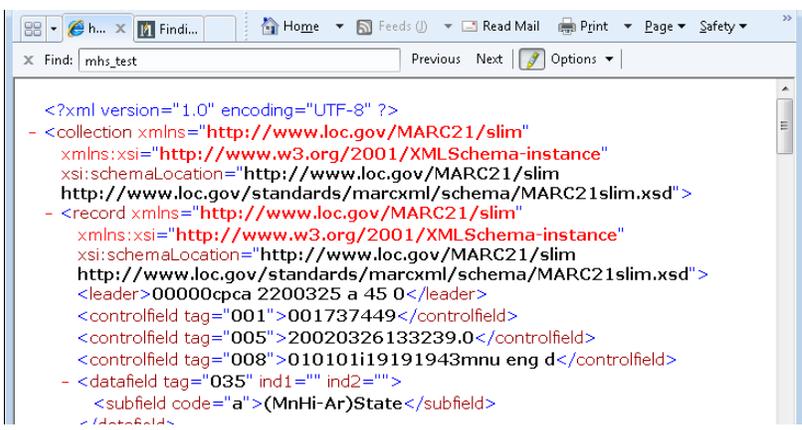
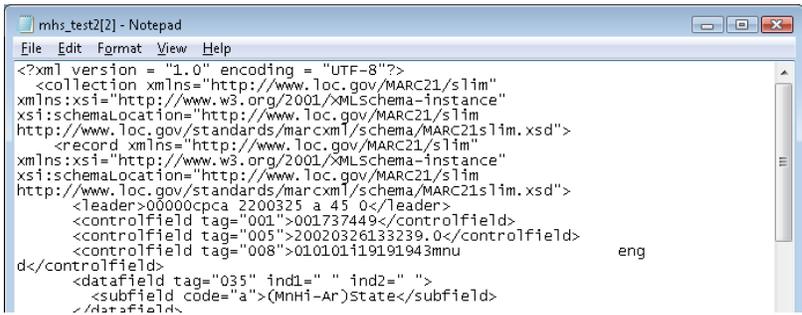
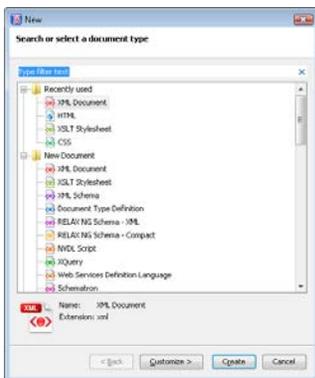
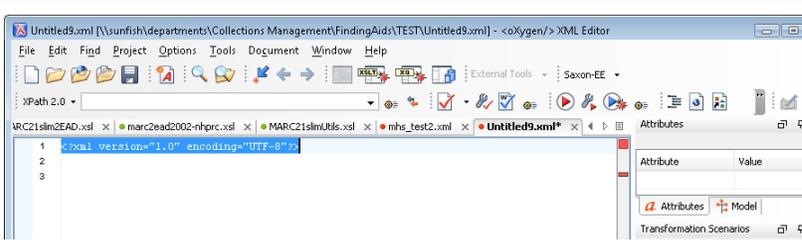
Find the file you downloaded and open it.



4. Copy and Paste the Downloaded File into Oxygen

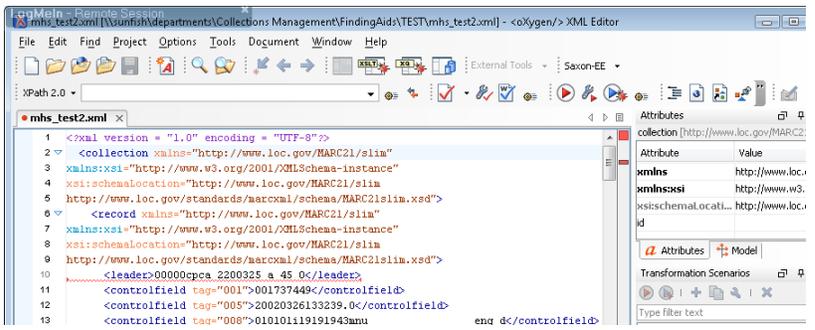
In Firefox, Select all the text (*Ctrl+a*), open Oxygen, and paste the copied text into a new XML document.

Internet Explorer will add collapsing markers that cannot be copied into Oxygen. When using IE, open the downloaded file, select *View* → *Source*, and copy the Notepad content, open Oxygen, and paste your clipboard into a new XML document.

<p>Open the file in Internet Explorer.</p>	
<p>In IE, select <i>View</i> → <i>Source</i> to open the file in Notepad.</p>	
<p>Open a new document in Oxygen by pressing <i>Ctrl+n</i> or by clicking the new document button.</p>	
<p>Select <i>XML Document</i>.</p>	
<p>Notice that Oxygen added an xml declaration to the new document.</p> <p>Select the declaration and paste your MARCXML copy over it.</p>	

Oxygen will report errors in the MARCXML file the <leader> and <controlfield tag="002">.

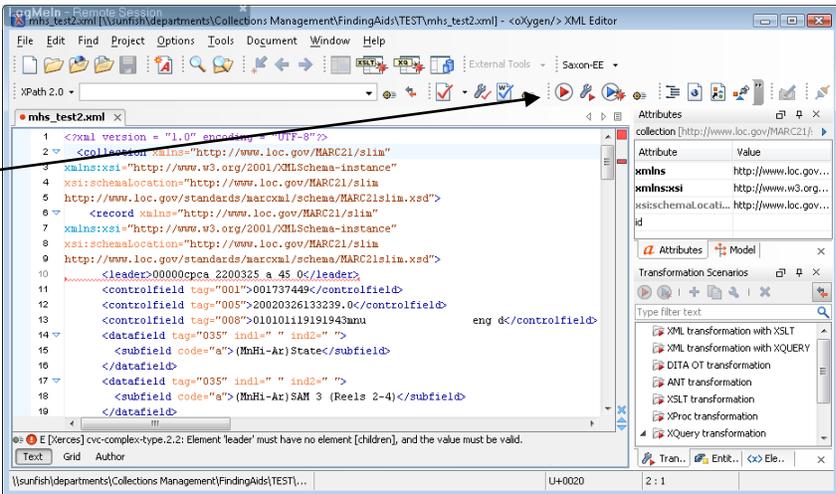
You can ignore these errors for the MARC2EAD scenario.



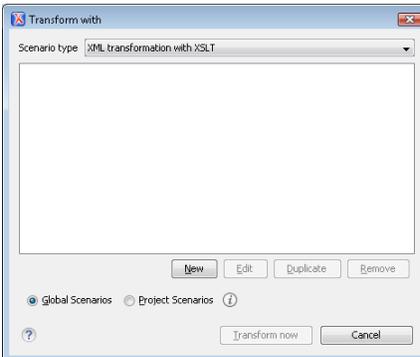
5. Transform MARC XML to EAD XML

The very first time you use this process, you will need to create a new transformation scenario by completing a set of dialog boxes. Thereafter, Oxygen will remember this scenario unless you remove it.

With the MARC XML file open in Oxygen, click on the *Apply Transformation Scenario* button  or press *Ctrl+Shift+T*



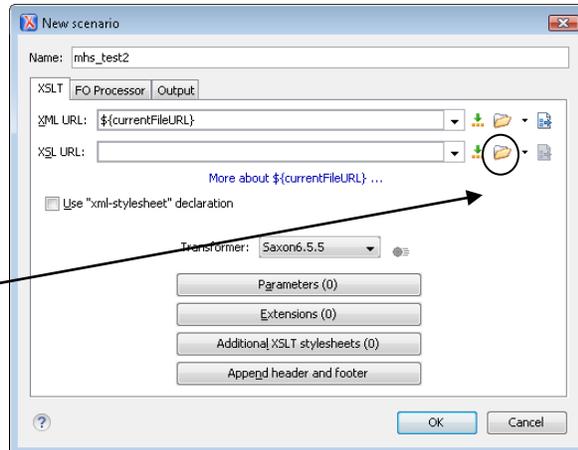
Click *New* in the *Transform With* box.



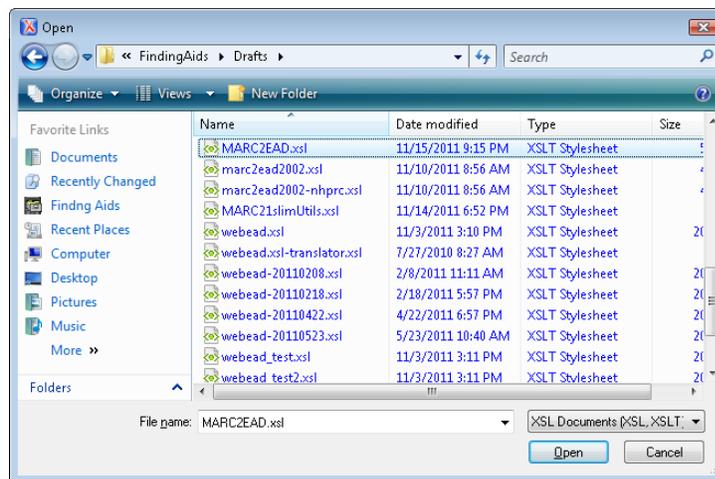
Oxygen will add the name of your EAD document to the *Name* field.

Leave the *XML URL* field as is: `$(currentFileURL)`

Click the file browse button  in the *XSL URL* field.

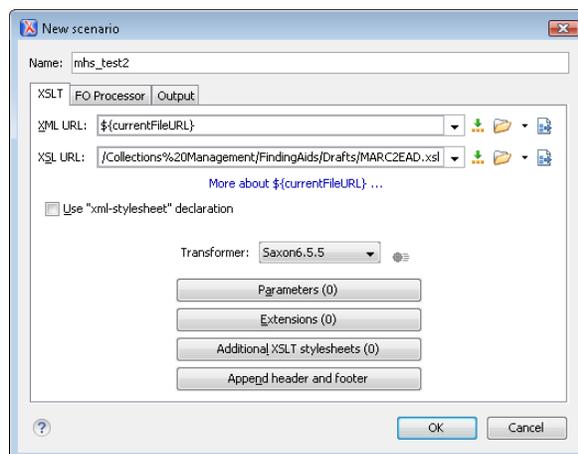


Browse to the *FindingAids/Drafts* directory and open *MARC2EAD.xml*.

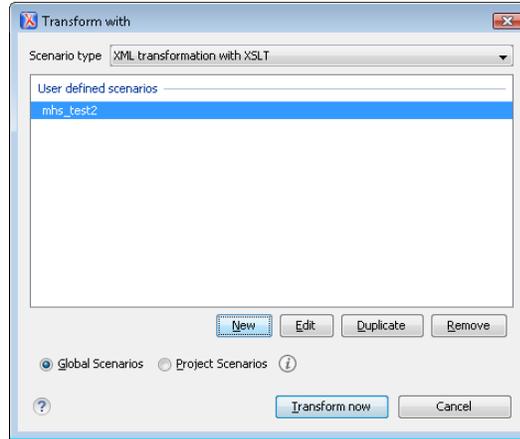


Oxygen adds the *MARC2EAD.xml* file in the *XSL URL* field.

Click *OK*.

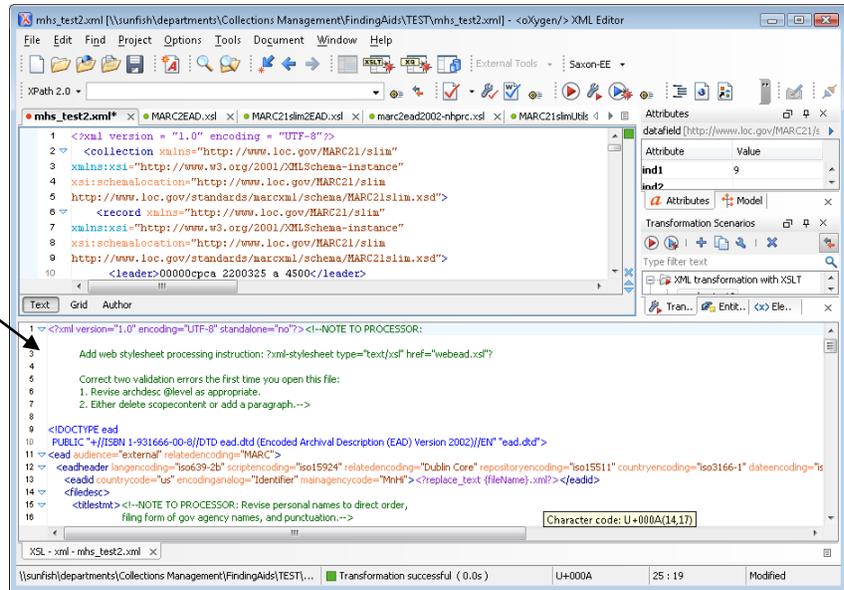


Click *Transform Now*.

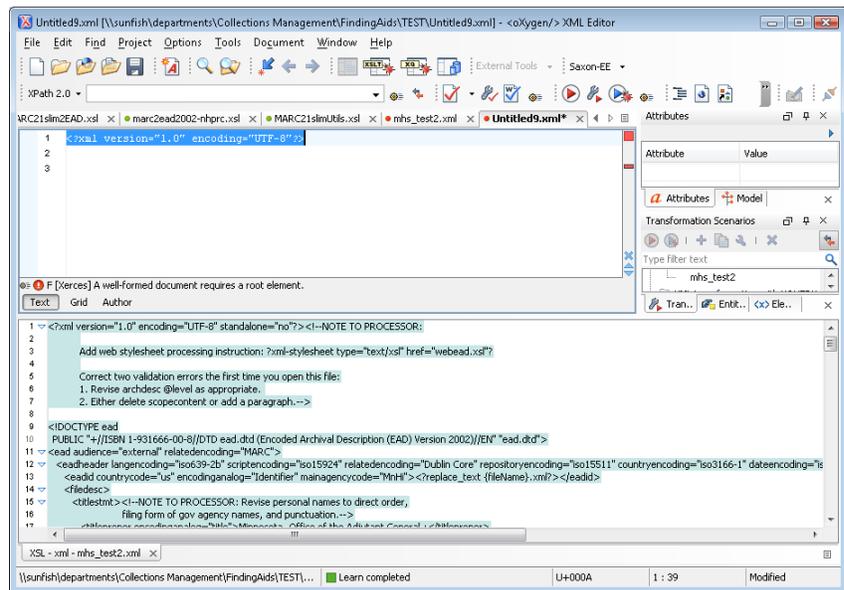


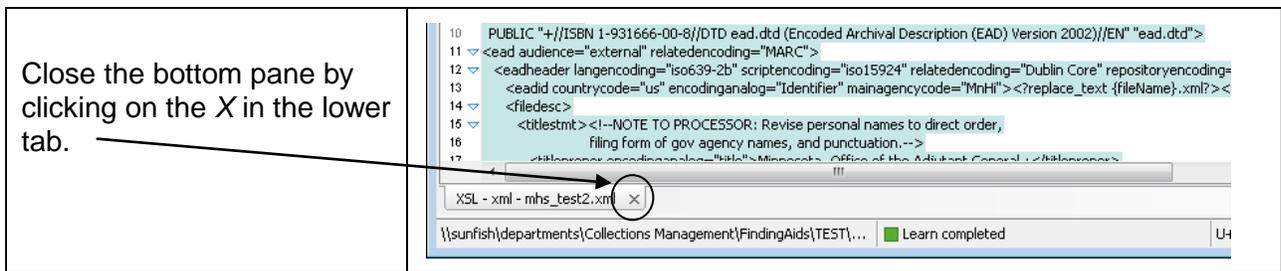
Oxygen will transform the MARC XML file into an EAD file that appears in a new pane at the bottom of the Oxygen window.

Place your cursor in the bottom pane. Copy the contents to your clipboard.



Open a new XML document and paste your copy over the new document's xml declaration.





After creating your new scenario, Oxygen will save the routine under you're the name of the first file you transformed.

Thereafter, to transform additional files, simply click . The *Transform With* dialog box will open. Select the name of your scenario and click *Transform Now*.

6. Edit the Resulting EAD XML File

Search for the strings "NOTE_TO_PROCESSOR" and "REPLACE_TEXT" to find attributes and other elements that you need to edit.

The very first note includes the web stylesheet processing instruction as well as reminders to edit two common validation errors. Edit this first note as shown below.

Edit all other notes as instructed in the EAD file. After editing each element, remember to delete the processor notes to keep the resulting XML files clean.

Transformed version to edit:	Edited version:
<pre><?xml version="1.0" encoding="UTF-8" standalone="no"?><!--NOTE_TO_PROCESSOR:</pre>	<pre><?xml version="1.0" encoding="UTF-8" standalone="no"?></pre>
<pre><?xml-stylesheet type="text/xsl" href="webead.xsl"?></pre> <p>Correct validation errors:</p> <ol style="list-style-type: none"> 1. Revise archdesc @level as appropriate. 2. Either delete scopecontent or add a paragraph.--> 	<pre><?xml-stylesheet type="text/xsl" href="webead.xsl"?></pre>

Legacy MARC Records: Add Aleph System Number to Finding Aid

For bibliographic records that were converted from old PALS into Aleph and that do not include a 001 field (records created before the 2004 migration), catalogers will need to type the Aleph System Number as the Catalog ID Number into the <processinfo><p>

element of the finding aid. New records that are exported from OCLC into Aleph or that have been reloaded in Aleph will include the 001 field and the MARC2EAD stylesheet will place the Aleph system number in the appropriate EAD tags.

Technical Note

Two MARC to EAD transformation stylesheets are stored in the FindingAids/Drafts directory. The first is the MHS default stylesheet. The second contains additional elements and coding for our NHPRC project.

1. MARC2EAD.xsl
2. MARC2EAD_nhprc.xsl

One additional xsl file, a MARC conversion utility set (*MARC21slimUtils.xsl*),² downloaded from the Library of Congress and also stored in our FindingAids/Drafts directory is called by our MARC2EAD procedure.

² Library of Congress, MARC 21 XML Schema: <http://www.loc.gov/standards/marcxml/>