

**EAD to MARC conversion
Version 200803**

**Mark Carlson
carlsonm@u.washington.edu
Special Collections Computer Support Analyst
University of Washington Libraries**

Installation

1. Unzip the EAD2MARC.zip archive to your PC. On Windows XP, you can unzip the archive by right-clicking on the EAD2MARC.zip file and choosing "Extract all..." Make a note of the location that you selected so that you can easily find the folder. This should create a folder called "EAD2MARC".
2. Open the EAD2MARC folder and then open the subfolder "bin" and double-click the "MarcEdit4_6a.exe" file to install MarcEdit version 4.6. My conversion script currently only works with MarcEdit version 4.6, so if you have a later version, you must still install this version.
3. Within the EAD2MARC folder, you should also see a script file called "ConvertEAD2MARC.wsf". Place any EAD XML encoded files within the EAD2MARC folder that you want to convert to MARC and then double-click the "ConvertEAD2MARC.wsf" script to run the conversion. You can convert multiple files at the same time. You should receive a "Done!" popup box after the conversion is complete.
4. A file with a ".dat" extension is created for each of the XML files after conversion. These .dat files are in MARC Communications format and can be imported into a MARC client such as OCLC Connexion client. See below under "Importing your .dat files into OCLC Connexion client".
5. There is additional information in the README.pdf file within the "bin" subfolder regarding other features that can be enabled.

About this stylesheet

The stylesheet (EAD2MARCXML.xsl) will convert an EAD version 2002 finding aid into MARCXML format based on the EAD to MARC21 crosswalk found at:

http://www.loc.gov/ead/tglib/appendix_a.html#a3

In addition, there are two user configurable files called "marcextras.xml" and "config.xsl" and a MARC Utility file (MARC21slimUtils.xml). These files must be placed in the same directory as the EAD2MARCXML.xsl file.

Importing your .dat files into OCLC Connexion client

1. Open OCLC Connexion client (it is not necessary to log in unless you want to save your imports to the online save file).

2. Click on File—Import records
3. If you have not logged on, choose “Import to local save file” – or log on to be able to import to online save file.
4. Click on “Browse” to navigate to the EAD2MARC folder where your newly converted .dat files are (or to whatever folder you may have moved them to if you moved them out of this folder)
5. Select the file that you want to import
6. Select whether you want to delete the original .dat file.
7. You should receive a confirmation “Successful imports (1)” Bibliographic records imported: 1
8. Click on Cataloging—Search—Local save file (or Online Save File if you chose that option) and click OK to open the browse screen. Your file should appear in the list.

User configuration (the “config.xml” file)

The config.xml currently contains 3 user settable parameters:

`includeMARCEXtrasFile`

This parameter tells the stylesheet whether to process the marceXtras.xml file during conversion. The marceXtras.xml file is used to import additional static MARC fields during conversion. To turn this feature on, change the default setting from “no” to “yes”. For more details about the marceXtras.xml file, see the information below.

`includeMARC856`

Many institutions use the MARC 856 field as a linking field to the finding aid on the WWW. If you want the stylesheet to automatically generate a MARC 856 field, set this parameter to “yes”.

`MARC856text`

This is a user customizable parameter that allows you to set your own preferred text for the MARC 856 field if you decide to include it.

Including static MARC fields (the “marceXtras.xml” file)

There may be local fields like MARC 852 or MARC 910 that you always want to include in every MARC record. You can place any additional MARC fields into the “marceXtras.xml” file and they will be imported during the conversion from EAD to MARC. You must set the parameter “includeMARCEXtrasFile” in the “config.xml” to “yes” for the stylesheet to import these additional fields. The imported fields **must** follow this format:

1. Subfields embedded within the text. The subfield indicators are indicated in the "ind1" and "ind2" attributes as below. The "fullstop" attribute tells the stylesheet whether to add a fullstop to the end of the field when appropriate. Implicitly encoding of subfield "a" is optional:

```
<MARC852 ind1="1" ind2="2" fullstop="no">
University of Washington Libraries$bSpecial Collections$eBox 352900, Seattle,
WA, 98195-2900
</MARC852>
```

2. In MARCXML syntax. In this case, the entire contents of the new field are inserted as is into the stylesheet. If you want a fullstop, you need to encode it in the data for this field. Be sure to follow this format exactly:

```
<MARC910>
<marc:datafield tag="910" ind1="" ind2="">
<marc:subfield code="a">MRC</marc:subfield>
</marc:datafield>
</MARC910>
```

MARC 100 110 111

In the high-level <archdesc> <did> <origination>, the stylesheet can generate the appropriate MARC fields as long as the @encodinganalog attribute is specified in either the <origination> element, or if there are child <persname>, <famname>, <corpname> elements. Since @encodinganalog "100" can map to either <persname> or <famname> elements in MARC, it is recommend that you use the appropriate child element under <origination> to more precisely indicate the correct MARC mapping.

For <corpname>, the stylesheet defaults to MARC 110 when encodinganalog is blank or not encoded as 111.

MARC 246

The stylesheet generates a MARC 246 field automatically. If the collection title contains the word "Papers" or "Records", then this text is used to generate subfield "a". Otherwise, subfield "a" is left blank and must be filled in manually.

MARC 254 255 256

Since <materialspec> is mapped to several MARC fields in EAD, the encodinganalog attribute must be mapped for the stylesheet to be able to determine which MARC field to map it to. If the encodinganalog attribute is not specified, then this MARC field is ignored

MARC 300

Formatting of the MARC 300 fields is a little tricky since there are many ways to use the EAD elements. This stylesheet assumes that there will be no more than 2 <extent> elements and that the second extent element should be formatted with parentheses around it. If this does not reflect your current usage of the extent element, you may need to alter this section to fit your needs.

MARC 340/538

Since <phystech> can map to either MARC 340 or 538, the correct mapping must be specified in the encodinganalog attribute.

MARC 610/611

Since <corpname> can map to either 610 or 611, if the MARC mapping is not specified in the encodinganalog attribute, then MARC 610 is assumed.

Controlled terms (MARC 1xx, 6xx, 7xx)

This stylesheet can process embedded subfield delimited subdivisions in controlaccess terms. To use this feature, use the \$ as a subfield delimiter along with the correct subfield code. E.g.:

```
<subject source="lcs" encodinganalog="650" rules="scm">Legislators$zWashington  
(State)$vArchives</subject>
```

```
<persname encodinganalog="600" source="lcnaf" role="subject">McGilvra, John  
J.$q(John Jay),$d1827-1903$vArchives</persname>
```

etc.

The use of subfield delimiters is optional. If you choose not to use them, the entire contents of the element will be placed in subfield "a" and you'll need to insert the subcodes manually.

Per the EAD crosswalk, <persname>, <corpname>, etc. map to 6xx when the encodinganalog attribute is 6xx or the role attribute is "subject".