

Mapping EndNote data to MARCXML

1 About this document

Author

Corey Wallis, RUBRIC Technical Officer

Purpose

This technical report outlines the way in which the EndNote to Fedora data migration scripts map fields in the EndNote data to fields in the MARCXML data.

Audience

RUBRIC Project Partners and other users of EndNote and Fedora

Requirements

An understanding of the structure of EndNote data represented as XML

An understanding of the structure of MARC tags and how they're represented in MARCXML

References

RUBRIC Technical Report: Migrating EndNote data to Fedora

Official EndNote website

<http://www.endnote.com/>

Official ARROW Project website

<http://www.arrow.edu.au/>

Official MARCXML website at the Library of Congress

<http://www.loc.gov/standards/marcxml/>

Notes

2 Background Information

In preparation for the EndNote to Fedora data migration it was necessary to map the available data in the EndNote XML file to the fields specified by the MARCXML templates provided by the ARROW project.

The map outlined below was developed using sample data provided by the University of Newcastle. The map was then used to develop the XSL transformation used by the Python script as part of the EndNote to Fedora data migration.

It should also be noted that the map created for patent items is not based on an ARROW template. At the time the data migration was written the ARROW project had not yet decided on common rules for storing metadata about patents. When an official template is produced this map may need adjustment.

3 EndNote to MARCXML data map

3.1 Data map definitions

The EndNote data to MARCXML data map contains the following columns:

MARC field

The MARC field as specified by the ARROW template

MARC Subfield

The MARC subfield as specified by the ARROW template

Description

The description of the field

EndNote XPath

The XPath to the field in the EndNote XML data. Each record is stored as a child of a record node, therefore the XPath as listed here is relative to the following node:

`/xml/records/record`

3.2 Common data mapping

MARC Field	MARC Subfield	Description	EndNote XPath
245			
	a	Item Title	/titles/title/style
100			
	a	Author Name	contributors/authors/author/style
	u	Author's organisation	<p>If the author is in the bold style use the institution value from the static-information.xml file.</p> <p>If the author is not in the bold style, do not output this tag</p> <p>See section 3.8 for details</p>
700			
	a	Author Name	contributors/authors/author/style
	u	Author's Organisation	<p>If the author is in the bold style use the institution value from the static-information.xml file.</p> <p>If the author is not in the bold style, do not output this tag</p> <p>See section 3.8 for details</p>
710			
	a	University Name	See section 3.8 for details
	b	Faculty / School / Dept	Need an additional set of data
520			
	a	Abstract	/abstract/style
300			
	a	Total number of pages	/pages/style
260			
	a	Place of publication	/pub-location/style
	b	Publisher	/publisher/style
	c	Year of publication	/dates/year
540			
	a	Copyright statement	See section 3.8 for details
856			

MARC Field	MARC Subfield	Description	EndNote XPath
	u	URL to publishers copy	/urls/related-urls/url
653			
	a	Keyword	/keywords/keyword/style
655			
	a	Item type	/ref-type[@name]
773			
	t	Collection Title	See section 3.8 for details
591			
	a	DEST collection year	See section 3.8 for details
592			
	a	DEST category	/custom1/style

3.3 Journal Article Specific Mapping

A journal article is defined as an EndNote record having the following ref-type node:

```
<ref-type name="Journal Article">17</ref-type>
```

MARC Field	MARC Subfield	Description	EndNote XPath
787			
	t	Journal Title	/periodical/full-title/style
	g	Citation	See section 3.6 for details
	n	Refereed Article	if /custom1/style is C1 the item is refereed Value is 1 for refereed, 0 for not refereed
022			
	a	ISSN	/isbn/style

3.4 Conference Proceeding Specific Mapping

A conference proceeding is defined as an EndNote record with the following ref-type node:

```
<ref-type name="Conference Proceedings">10</ref-type>
```

MARC Field	MARC Subfield	Description	EndNote XPath
711			
	a	Conference Name	/titles/tertiary-title
	c	Conference Location	/custom2/style
	d	Conference Date	/custom3/style
787			
	g	Citation	See section 3.7 for details
	n	Refereed	if /custom1/style is E1 the item is refereed Value is 1 for refereed, 0 for not refereed
020			
	a	ISBN	/isbn/style

3.5 Patent Specific Mapping

A patent is defined as an EndNote record with the following ref-type node:

```
<ref-type name="Patent">25</ref-type>
```

MARC Field	MARC Subfield	Description	EndNote XPath
013			
	\$a	Patent Number	/number/style
653			
	\$a	Keywords (additional patent numbers)	/isbn/style (will need to break up based on semi colon and add one tag per item)

3.6 Building the Citation Field for Journal Articles

The ARROW MARCXML templates specify that field 787, subfield g, is used to store a citation. The EndNote data has components of a citation in different fields. The following map outlined the components of the citation, and how they are concatenated to create a complete citation.

Variable	Part of the citation	EndNote XPath
[vol]	Volume	/volume/style
[no]	Number	/number/style
[pub-date-1]	Publication Date	/dates/pub-dates/date
[pub-date-2]	Publication Year	/dates/pub-dates/year
[pages]	Pages	/pages/style

The citation is compiled as follows to adhere to the style provided by ARROW.

Vol. [vol], no. [no] ([pub-date-1]. [pub-date-2]), p. [pages]

3.7 Building the Citation Field for Conference Proceedings

The ARROW MARCXML templates specify that field 787, subfield g, is used to store a citation. The EndNote data has components of a citation in different fields. The following map outlined the components of the citation, and how they are concatenated to create a complete citation.

Variable	Part of the citation	EndNote XPath
[vol]	Volume	/volume/style
[pub-date]	Publication Year	/dates/pub-dates/year/style
[pages]	Pages	/pages/style

The citation is compiled as follows.

[vol], ([pub-date-2]), p. [pages]

3.8 Static Information

There are a number of fields specified by the MARCXML templates that require data that is not stored in the EndNote XML file. These values are stored in the **static-information.xml** file and are retrieved during the XSL transformation. These fields are outlined below, instructions for editing the **static-information.xml** file are included in RUBRIC Technical Report: Migrating EndNote data to Fedora.

MARC Field	MARC Subfield	Description	Static-information XPath
100			
	u	Author's Organisation	/institution
700			
	u	Author's Organisation	/institution
710			
	a	Name of University	/institution
540			
	a	Copyright statement	/copyright-statement
773			
	t	Repository Collection Title	/repository-title
591			
	a	DEST collection year	/dest-collection-year