Enabling OAI & Mapping Fields in Digital Commons

Produced by the Digital Collections Working Group of the Content Creation & Dissemination Team
Janet Hauck, Whitworth University (chair)
Anneliese Dehner, Metadata Applications Librarian, Alliance (author of this document)
Laura Zeigen, Oregon Health & Science University
Julia Simic, University of Oregon
Theodore Gerontakos, University of Washington
Jodi Allison-Bunnell, Alliance (ex officio)

Version 1.1, June 2017

Table of Contents

Overview
Enable OAI
Review your OAI output
Questions to ask as you review your OAI output
Review your current field mapping for a collection
Update your field mapping for a collection
  Add or Edit Dublin Core mapping for an already-existing Digital Commons field
  Disable a Digital Commons field's OAI output
  Reorder fields in a collection's OAI output

Overview
Mapping metadata fields to comply with the Alliance Dublin Core Best Practices Guidelines (DCBPG) will help to create consistently structured records in aggregation. This documentation explains how to enable OAI in Digital Commons and how to map Digital Commons fields to Dublin Core elements for generation of the desired OAI output.
Enable OAI

To configure your OAI settings, first enable OAI in Digital Commons by contacting your Digital Commons consultant. For more information, refer to Digital Commons's OAI Support documentation.

Review your OAI output

Review a set’s metadata with the Dublin Core Mapping Checker.

Provide a valid OAI-PMH base URL, then choose a set to get a table of Simple Dublin Core fields, arranged in columns. “Required” fields are listed first, followed by the “Recommended” and “Optional” fields. The table’s columns list all record values for a given field, rather than ordering the content in record-specific rows. The intention is to display the data mapped to each field, to catch mapping errors, and to identify data for cleanup in the local repository.

As you review your OAI output, you may discover problems in the way your Digital Commons fields are mapped to Dublin Core elements. The questions below will help you identify mapping problems.

Questions to ask as you review your OAI output

1. Are any fields missing that you expect to be present in the OAI output (eg. fields with Custom Export Labels)?

2. Do you see data in the OAI output that shouldn’t be there?

3. Is your data present but in the wrong field?

4. Is the wrong date present in the OAI output?

5. Are multiple date fields present in the OAI output?

If you answered yes to any of these questions, you probably need to edit the way your Digital Commons fields are mapped to Dublin Core elements. To understand your OAI output, review your current field mapping for a collection.

Review your current field mapping for a collection

To view a collection’s field mapping, look at the set’s OAI output in a web browser. Use the pattern below to form an OAI request. In your request, set the metadataPrefix to simple-dublin-core, as this is how the Alliance harvester will harvest your data.

http://[Site Url]/do/oai/?verb=ListRecords&metadataPrefix=simple-dublin-core&set=[Series/Journal/Other URL Label here]

Your output will list the first 100 records for the specified collection/set. To see the next page of 100
records, you will need to use a resumption token in your OAI request. The resumption token is found at
the bottom of the first page of your OAI output, and it looks like this:

```xml
<resumptionToken completeListSize="427" cursor="0">3111673/simple-dublin-core/100//</resumptionToken>
```

To see the next page of records in your OAI output, copy this resumption token (eg. 
3111673/simple-dublin-core/100/) and paste it into your OAI request. Use the pattern below to
form an OAI request that includes a resumption token:

http://[Site Url]/do/oai/?verb=ListRecords&resumptionToken= [Paste the resumption token here]

Your OAI output will list records that look like the record below. To understand how your fields are
mapped, compare your OAI output to the field configuration in your Digital Commons administrative
interface.

```
<record>
  <header>
    <identifier>oai:docs.lib.purdue.edu:jtrp-1851</identifier>
    <datestamp>2011-07-12T22:01:15Z</datestamp>
    <setSpec>publication:jtrp</setSpec>
  </header>
  <metadata>
    <oai_dc:dc xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
      xmlns:dc="http://purl.org/dc/elements/1.1/"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
        http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
      <dc:title>Driver Obedience to Stop and Slow Signs</dc:title>
      <dc:creator>Jackman, William Thomas</dc:creator>
      <dc:type>Technical Report</dc:type>
      <dc:identifier>FHWA/IN/JHRP-56-08</dc:identifier>
      <dc:date>1956-01-01T08:00:00Z</dc:date>
      <dc:subject>Civil Engineering</dc:subject>
      <dc:identifier>http://docs.lib.purdue.edu/jtrp/362</dc:identifier>
    </oai_dc:dc>
  </metadata>
</record>
```

Update your field mapping for a collection

In reviewing your OAI output, you may have identified some mapping problems. Those problems can be
solved by contacting your Digital Commons consultant and requesting that he/she updates your Dublin
Core field mapping. Updates may include adding, editing, or removing the Dublin Core elements mapped
to your Digital Commons fields, or changing the order of fields in a record.
Add or Edit Dublin Core mapping for an already-existing Digital Commons field

When to add Dublin Core mapping to an already-existing field

- When you don’t see data from a Digital Commons field in your OAI output, but you want it to be harvested.

When to edit an already-existing field’s Dublin Core mapping

- When data from a Digital Commons field is present in your OAI output, but it’s mapped to the wrong Dublin Core element.

How to add or edit an already-existing field’s Dublin Core mapping

1. Identify the Digital Commons field to map.
2. Select the appropriate Dublin Core element to map to your Digital Commons field. Your Digital Commons fields must be mapped to Dublin Core fields to enable output using OAI-PMH.
   
   For more information about the guidelines for these fields, refer to [DCBP]. See the Additional Information below for info about mapping to Qualified Dublin Core fields.

3. Contact your Digital Commons consultant and request the Dublin Core mapping change for simple-dublin-core output. Specify whether you would like to make this change for a specific collection or for your entire repository.

Additional information

It’s not necessary to follow a strict 1:1 correspondence between Digital Commons metadata fields and Dublin Core elements. You can map several Digital Commons fields to the same Dublin Core elements, if this is relevant to your collection. For example, DCBP recommends mapping both “Genre” and “Type” values to the Dublin Core element “Type.”

Please note that the harvester conforms to the oai_dc schema as recommended by the OAI-PMH 2.0 specification, meaning all fields mapped to Qualified Dublin Core will output as Simple Dublin Core. This is possible because all Dublin Core fields that we call “Qualified” are refinements of the Simple Dublin Core fields; the Qualified fields can therefore be reduced to the elements they refine. For example, the Qualified fields “Relation-Is Part Of” and “Relation-References” reduce to the Simple field they refine, “Relation.”

The Alliance harvester will harvest only the following 14 Simple Dublin Core fields:

<table>
<thead>
<tr>
<th>Field</th>
<th>Field</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contributors</td>
<td>Format</td>
<td>Source</td>
</tr>
<tr>
<td>Coverage</td>
<td>Identifier</td>
<td>Subject</td>
</tr>
<tr>
<td>Creator</td>
<td>Language</td>
<td>Title</td>
</tr>
<tr>
<td>Date</td>
<td>Relation</td>
<td>Type</td>
</tr>
<tr>
<td>Description</td>
<td>Rights</td>
<td></td>
</tr>
</tbody>
</table>

The reduction to Simple Dublin Core has two known issues for Alliance-harvested metadata:
- **Relation**: A collection’s finding aid URL (or ARK) in Archives West is the only value that will be harvested from the Relation Dublin Core fields. All other values in “Relation”, “Relation-Is Part Of”, “Relation-References”, etc. will be ignored.

- **Coverage**: “Temporal Coverage” and “Spatial Coverage” will both output as “Coverage”. The use of “Temporal Coverage” is discouraged by DCBPG, and mapping to this element should be avoided.

**Disable a Digital Commons field’s OAI output**

When to disable a field’s OAI output

- When you see a field’s data in the OAI output, and you don’t want it to be harvested.

- When multiple Qualified Dublin Core fields are output as the same Simple Dublin Core element, and the resulting output is not compliant with DCBPG. (eg. “Temporal Coverage” and “Spatial Coverage” will both output as “Coverage” -- in which case Dublin Core mapping of the field for “Temporal Coverage” should be removed).

How to disable a field’s OAI output

1. Identify the Digital Commons field to disable.

2. Contact your Digital Commons consultant and request that this Dublin Core mapping change for simple-dublin-core output. Specify whether you would like to make this change for a specific collection or for your entire repository.
When to reorder

- When you have multiple date fields, they are all mapped to the Dublin Core element “Date”, and the compliant date for harvest is not positioned first among the date fields. In the output above, for example, the 1956 date would be harvested by the Alliance harvester. But the 1958 date would not.

How to reorder

1. Identify the Digital Commons field(s) to reorder.

2. Contact your Digital Commons consultant and request this change for a specific collection or your entire repository.

Additional information

Reordering fields will affect the order in which fields display in both Digital Commons Administration and the public interface.