Using Open Source Tools to expose YCBA Collection Data in the LIDO schema for OAI-PMH Harvesting
Cultural institution websites
Digital Public Library of America
Google Art Project
ArtStor
Europeana
Yale University Cross Collection Discovery
eMuseum Network
Other data aggregators
Bare Bones Data Export Tool
Quest for the Holy Grail

Open Source Tools
Easy to use, Inexpensive, low entry barrier
Museum Data Exchange Project (OCLC)

COBOAT
CDWA Lite
OAICATMuseum
OAI-PMH

CDWA Lite – standard schema in 2009

OCLC releases software suite to help museums exchange data

DUBLIN, Ohio, USA, 22 May 2009—OCLC Research has released a software suite to help museums exchange object descriptions and share data, the result of a cooperative effort made possible by a grant from the Andrew W. Mellon Foundation to further develop infrastructure for museum data exchange.

Museums now have access to COBOAT and OAICATMuseum 1.0 software. COBOAT is a metadata publishing tool developed by Cognitive Applications Inc. (Cogapp) that transfers information between databases (such as collections management systems) and different formats. As configured for this project, COBOAT allows museums to extract standards-based records in the Categories for the Descriptions of Works of Art (CDWA) Lite XML data format out of Gallery Systems TMS, the museum industry’s leading collection management system. Configuration files allow COBOAT to be adjusted for extraction from different vendor-based or homegrown database systems, or locally divergent implementations of the same collections management system.

COBOAT software is now available on the OCLC Web site under a fee-free license for the purposes of publishing a CDWA Lite repository of collections information at www.oclc.org/research/activities/COBOAT/default.htm.

OAICATMuseum 1.0 is an Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) data content provider supporting CDWA Lite XML. It allows museums to share the data extracted with COBOAT using OAI-PMH. OAICATMuseum was developed by OCLC Research and is available under an open source license online at www.oclc.org/research/activities/oaicatmuseum/default.htm.
CDWA Lite – standard schema in 2009

LIDO – standard schema in 2010
Cogapp has been working in the field of museum collection data for over twenty years, and has long supported efforts to facilitate cross-museum searching and the exchange of metadata between museums. In support of this, we are delighted to renew our offer of no-cost licenses to museums wishing to use COBOAT to publish a LIDO repository of information on their collections, as we previously did in relation to CDWA Lite.

Any institution with an existing license to use COBOAT to publish CDWA Lite data is hereby authorized to use it additionally or instead to generate LIDO data for publication.

Ben Rubinstein  benr@cogapp.com
COBOAT

by modifications originally made to OAICat by the Getty, and has been uniquely tailored to the needs of museums as a result of the Museum Data Exchange Project.

An updated version 1.1 of OAICatMuseum is also available, including support for identifying deleted records and the addition of LIDO XML output. This development was prompted by the use of OAICatMuseum and COBOAT by the Yale Center for British Art, one of the first US museums to export data in the LIDO schema.

License

This software may be used without charge in accord with the terms of the Apache License, Version 2.0.

Software

- OAICatMuseum v1.0 War File and Documentation (txt: 8K/4,403 char.)
- OAICatMuseum v1.1 War File and Documentation (txt: 9K/4,787 char.)
- oaiatdb.sql — A MySQL import file which is referenced in the README.txt file and can be used to test OAICatMuseum

Questions regarding the LIDO schema, supported by OAICatMuseum v1.1, can be directed to David Parsell, Systems Manager Collections Information and Access, Yale Center for British Art [david.parsell@yale.edu].
COBOAT Overview

- Catalog collection
- Harvesters Yale CCD
- Search YCBA collection on web site

MySQL
OAICAT
Apache Tomcat
COBOAT → collection data export

- Script 1: TMSretrieve
- Script 2: arrays
- Script 3: template
- Compiled -> build xml doc.

Tools:
- Notepad++
- Oxygen
- MySQL
- OAICATMuseum
COBOAT xml document template → validate in Oxygen

Problems

well formed
Yale Center for British Art

Compiled functions

Export data
Build xml
documents

Incremental export to
OAICATMuseum

Delete released object records
COBOAT delete function, send blank xml document to OAIcatmuseum

<?xml version="1.0" encoding="UTF-8"?>
- <OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
                       http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
    <responseDate>2012-04-27T12:44:05Z</responseDate>
    <request identifier="oai:1437tms ycba.yale.edu:1437" metadataPrefix="lido"
        verb="GetRecord">http://bac2.yu.yale.edu:8080/oaicatmuseumlido/OAIHandler</request>
    - <GetRecord>
        - <record>
            - <header status="deleted">
                <identifier>oai:tms ycba.yale.edu:1437</identifier>
                <datestamp>2011-12-07</datestamp>
                <setSpec>ycba:ps</setSpec>
            </header>
        </record>
    </GetRecord>
</OAI-PMH>
Yale Center for British Art

Collection server
COBOAT

Web server
MySQL DB
OAICATMuseum
Apache Tomcat

Harvesters
xml documents

Firewall
<table>
<thead>
<tr>
<th>Creator</th>
<th>Joseph Mallord William Turner, 1775-1851, British</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Dort or Dordrecht: The Dort Packet-Boat from Rotterdam Becalmed</td>
</tr>
<tr>
<td>Date</td>
<td>1818</td>
</tr>
<tr>
<td>Medium</td>
<td>Oil on canvas</td>
</tr>
<tr>
<td>Dimensions</td>
<td>62 x 92 inches (157.5 x 233.7 cm) Frame: 73 x 102 x 5 inches (185.4 x 259.1 x 12.7 cm)</td>
</tr>
<tr>
<td>Inscription(s)/</td>
<td>Inscribed, lower right: &quot;Dort&quot;, Signed and dated, lower right: &quot;JMW Turner RA 1818&quot;</td>
</tr>
<tr>
<td>Marks/</td>
<td></td>
</tr>
<tr>
<td>Lettering</td>
<td></td>
</tr>
<tr>
<td>Credit Line</td>
<td>Yale Center for British Art, Paul Mellon Collection</td>
</tr>
<tr>
<td>Accession</td>
<td>B1977.14.77</td>
</tr>
<tr>
<td>Number</td>
<td></td>
</tr>
<tr>
<td>Collection</td>
<td>Paintings and Sculpture</td>
</tr>
<tr>
<td>Link to This</td>
<td><a href="http://collections.britishart.yale.edu/vufind/Record">http://collections.britishart.yale.edu/vufind/Record</a> /1667701</td>
</tr>
<tr>
<td>Record</td>
<td></td>
</tr>
<tr>
<td>Subject Terms</td>
<td>sea</td>
</tr>
<tr>
<td>Place Represented</td>
<td>Our Dear Lady Church</td>
</tr>
<tr>
<td>Publications</td>
<td>Julia Marciari-Alexander, This other Eden, paintings</td>
</tr>
</tbody>
</table>
URI: http://collection.britishart.yale.edu/id/object/34

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>?:P102_has_title</td>
<td><a href="http://collection.britishart.yale.edu/id/object/34/title/1">http://collection.britishart.yale.edu/id/object/34/title/1</a></td>
</tr>
<tr>
<td>?:P12i_was_present_at</td>
<td><a href="http://collection.britishart.yale.edu/id/exhibition/248">http://collection.britishart.yale.edu/id/exhibition/248</a></td>
</tr>
<tr>
<td>?:P12i_was_present_at</td>
<td><a href="http://collection.britishart.yale.edu/id/exhibition/605">http://collection.britishart.yale.edu/id/exhibition/605</a></td>
</tr>
<tr>
<td>?:P12i_was_present_at</td>
<td><a href="http://collection.britishart.yale.edu/id/exhibition/88">http://collection.britishart.yale.edu/id/exhibition/88</a></td>
</tr>
<tr>
<td>?:P1_is_identified_by</td>
<td><a href="http://collection.britishart.yale.edu/id/object/34/TMS">http://collection.britishart.yale.edu/id/object/34/TMS</a></td>
</tr>
<tr>
<td>?:P1_is_identified_by</td>
<td><a href="http://collection.britishart.yale.edu/id/object/34/ocd">http://collection.britishart.yale.edu/id/object/34/ocd</a></td>
</tr>
<tr>
<td>?:P1_is_identified_by</td>
<td><a href="http://collection.britishart.yale.edu/id/object/34/inventory-number">http://collection.britishart.yale.edu/id/object/34/inventory-number</a></td>
</tr>
<tr>
<td>?:P1_is_identified_by</td>
<td><a href="http://collection.britishart.yale.edu/id/object/34/lidoRecID">http://collection.britishart.yale.edu/id/object/34/lidoRecID</a></td>
</tr>
<tr>
<td>?:P2_has_type</td>
<td><a href="http://collection.britishart.yale.edu/id/thesauri/300033618">http://collection.britishart.yale.edu/id/thesauri/300033618</a></td>
</tr>
<tr>
<td>?:P43_has_dimension</td>
<td><a href="http://collection.britishart.yale.edu/id/object/34/height">http://collection.britishart.yale.edu/id/object/34/height</a></td>
</tr>
<tr>
<td>?:P43_has_dimension</td>
<td><a href="http://collection.britishart.yale.edu/id/object/34/width">http://collection.britishart.yale.edu/id/object/34/width</a></td>
</tr>
<tr>
<td>?:P50_has_current_keeper</td>
<td><a href="http://collection.britishart.yale.edu/id/thesauri/department">http://collection.britishart.yale.edu/id/thesauri/department</a></td>
</tr>
<tr>
<td>?:P55_has_current_location</td>
<td><a href="http://collection.britishart.yale.edu/id/object/34/location/1">http://collection.britishart.yale.edu/id/object/34/location/1</a></td>
</tr>
<tr>
<td></td>
<td><a href="http://collection.britishart.yale.edu/id/object/34/location/2">http://collection.britishart.yale.edu/id/object/34/location/2</a></td>
</tr>
<tr>
<td></td>
<td><a href="http://collection.britishart.yale.edu/id/object/34/location/3">http://collection.britishart.yale.edu/id/object/34/location/3</a></td>
</tr>
<tr>
<td></td>
<td><a href="http://collection.britishart.yale.edu/id/object/34/location/4">http://collection.britishart.yale.edu/id/object/34/location/4</a></td>
</tr>
</tbody>
</table>
Harvest XML Data

Harvesting XML data can be accomplished with these few simple steps and software tools. For more information about specific terminology please refer to the Become More Familiar section, Metadata Updates, and Metadata Schema Explanation. Access to or use of the Center's data and services is subject to the Center's Open Data and Data Services Terms of Use.

1. Get metadata by harvesting LIDO XML (via OAI-PMH)
   http://collections.britishart.yale.edu/oaicatmuseum/
2. Access images via URIs

Linked Open Data

Accessing collections via Linked Open Data can be accomplished using SPARQL Protocol and RDF Query Language (SPARQL). These simple steps will help with linking to our objects as well as getting at the data programmatically. Access to or use of the Center's data and services is subject to the Center's Open Data And Data Services Terms of Use.

1. Search for objects in our system via online collection or SNORQL interface
2. Reference our objects via URI in your data: http://collection.britishart.yale.edu/id/object/499
3. Test your SPARQL queries with our web form
4. Browse examples of our data using Pubby
5. Connect your application to our SPARQL endpoint: http://collection.britishart.yale.edu/openrdf-sesame/repositories/ycba

How the technology works:

Exposing collections as part of Linked Open Data allows available data to be provided via URIs, annotated using Ontologies, and linked with other institutions in the form of a knowledge graph. The Center's graph is in an evolving state and based on CRM ontology that is used to describe relationships between many components of the graph.
Harvest LIDO xml for CRM

- Cataloging TMS
- COBOAT
- OAICat
  MySQL
  Apache Tomcat

Harvest data - Yale CCD

Search YCBA collection on web site

CRM
Linked Open Data
Google Art Project launched April 2012

Yale Center For British Art

Whitlingham, Norfolk
1860
Frederick Sandys
Yale Center for British Art
Google Art Project launched April 2012

Thomas Rowlandson (Creator)

(1756 - 1827)
Nationality: British
emuseum network and ARTSTOR

Discover Collections from Around the World

ARTstor Digital Library

ARTstor
AN IMAGE LIBRARY FOR THE ARTS AND SCIENCES
Simple Tool = Leverage Web Exposure

david.parsell@yale.edu
britishart.yale.edu