

A practical way to get shared URLs into Cultural Heritage data

Richard Light

Richard Light Consultancy

Open Data on the Web, 24 April 2013

Cultural heritage data sharing

- IRGMA multidisciplinary data standard (1960s)
- MDA data standard (1980s)
- CIDOC Reconciliation of standards WG (1980s)
- CIDOC relational data model (1990s)
- SPECTRUM (1990s-) incl. XML Schema
- CIDOC Conceptual Reference Model (1990s-)

All structural frameworks

Terminology control

- Vocabularies/thesauri
 - Getty vocabs. (AAT, ULAN, TGN)
 - ICOM Costume vocabulary (1981)
 - Herts. Simple Name termlist
- Classifications
 - ICONCLASS
 - SHIC

Wide use; local customization; 'locked into' systems

Cultural Linked Data (1)

- Some frameworks: notably ICONCLASS

```
- <rdf:RDF>
  - <rdf:Description rdf:about="http://iconclass.org/52D2">
    <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
    <skos:prefLabel xml:lang="fr">Clarté, Précision</skos:prefLabel>
    <skos:prefLabel xml:lang="en">Clarity, Precision</skos:prefLabel>
    <skos:prefLabel xml:lang="cn">Clarity, Precision</skos:prefLabel>
    <skos:prefLabel xml:lang="de">Klarheit, Genauigkeit, Präzision</skos:prefLabel>
    <skos:prefLabel xml:lang="it">chiarezza, precisione</skos:prefLabel>
    <skos:prefLabel xml:lang="fi">Clarity, Precision</skos:prefLabel>
    <skos:inScheme rdf:resource="http://iconclass.org/rdf/2011/09"/>
    <skos:notation>52D2</skos:notation>
    <skos:narrower rdf:resource="http://iconclass.org/52D2%28%2B0%29"/>
    <skos:narrower rdf:resource="http://iconclass.org/52D2%28%2B1%29"/>
    <skos:narrower rdf:resource="http://iconclass.org/52D2%28%2B2%29"/>
    <skos:narrower rdf:resource="http://iconclass.org/52D2%28%2B3%29"/>
    <skos:narrower rdf:resource="http://iconclass.org/52D2%28%2B4%29"/>
    <skos:narrower rdf:resource="http://iconclass.org/52D2%28%2B5%29"/>
    <skos:narrower rdf:resource="http://iconclass.org/52D2%28%2B6%29"/>
    <skos:narrower rdf:resource="http://iconclass.org/52D21"/>
    <skos:broader rdf:resource="http://iconclass.org/52D"/>
  </rdf:Description>
</rdf:RDF>
```

Cultural Linked Data (2)

- Some resources, notably BM:

```
- <rdf:RDF>
- <rdf:Description rdf:about="http://collection.britishmuseum.org/id/object/EAF119772">
  <ns1:P24B.changed_ownership_through rdf:resource="http://collection.britishmuseum.org/id/object/EAF119772/acquisition"/>
  <ns1:P2F.has_type rdf:resource="http://collection.britishmuseum.org/id/thesauri/x5373"/>
  <ns1:P2F.has_type rdf:resource="http://collection.britishmuseum.org/id/thesauri/x7812"/>
  <ns1:P3F.has_note>Acquisition date :: 1994 :: </ns1:P3F.has_note>
  <ns1:P3F.has_note>Object type :: basket :: </ns1:P3F.has_note>
  <ns1:P3F.has_note>Object type :: marriage equipment :: </ns1:P3F.has_note>
  <ns1:P46F.is_composed_of rdf:resource="http://collection.britishmuseum.org/id/object/EAF119772/1"/>
  <ns1:P50F.has_current_keeper rdf:resource="http://collection.britishmuseum.org/id/department/E"/>
  <ns1:P52F.has_current_owner rdf:resource="http://collection.britishmuseum.org/id/the-british-museum"/>
  <ns1:P57F.has_number_of_parts>1</ns1:P57F.has_number_of_parts>
  <ns2:codex_id>3380108</ns2:codex_id>
  <ns2:reg_id>Af1994,14.1</ns2:reg_id>
  <rdf:type rdf:resource="http://collection.britishmuseum.org/id/crm/E22.Man-Made_Object"/>
  <ns3:sameAs rdf:resource="http://collection.britishmuseum.org/id/codex/3380108"/>
</rdf:Description>
</rdf:RDF>
```

Current state

- Cross-sectoral discussions (libraries, archives, museums)
- Design patterns e.g. uncertainty, attribution
- Potential use of CIDOC CRM as ‘mortar’ (awaiting BM guidelines)
- Few shared ontologies expressed as Linked Data
- Linked Data resources full of strings, or silos

Desired state

- Cultural heritage resources routinely published as Linked Data
- Intra-domain structural compatibility, e.g. using CIDOC CRM
- Shared use of central authorities for places, people, events, leading to ...
- Distributed database of cultural heritage/history

Modes and data control

- Modes: native XML database, widely used for cataloguing U.K. museums
- Supports “termlists”: primary keys in one file control data values in another
- Publishes a standard set of “termlists”
- Extended termlist concept to “web termlists”: use a web service to control data
 - a means of getting LD URLs into Modes data

Standard Modes termlist

BCRTM:M1930.1	watch	Pocket watch, silver	purchase	Windermere, 40 Stanpit,	16.4.1930
BCRTM:M1930.1a	postcard	Pocket watch, silver	purchase	Windermere, 40 Stanpit,	16.4.1930
BCRTM:M1930.1b	postcard	Pocket watch, silver	purchase	Windermere, 40 Stanpit,	16.4.1930
BCRTM:M1930.1d					
BCRTM:M1930.2					
BCRTM:M1930.3					
BCRTM:M1931.132.1					
BCRTM:M1931.132.2					
BCRTM:M1931.132.3					
BCRTM:M1934.23					
BCRTM:M1934.31					
BCRTM:M1935.2.1					
BCRTM:M1935.2.2					
BCRTM:M1935.2.3					
BCRTM:M1935.2.4					
BCRTM:M1935.2.5					
BCRTM:M1935.2.6					
BCRTM:M1935.106					

Select from termlist

HTML view | Edit view


Terms





- bequest
- brought
- brought-along
- brought-in
- brought-out
- brought-over
- commission
- copy
- deposit
- enquiry
- exchange
- filched

Selection

bequest

Prefix/search

Show non preferred terms  Update

 Select term  Create term  Cancel  Help

“Web termlist” (Geonames)

The screenshot shows a software interface with a data entry form on the left and a 'Web termlist' dialog box on the right. The data entry form includes fields for Classification (trade, 3.42), BriefDescription (Pocket watch, silver case, made by Andrew Dickie), Identification, Description (serial no 1772, Method engraved), Production (clockmaker Dickie, Andrew), Association (Period 19th century, late), Acquisition (Method purchase, Person from Ferry, Francis, Date 16.4.1930), and Disposal (Method stolen, Date 6.4.1998). The 'Web termlist' dialog box has a title bar 'm Select from termlist' and a 'Terms' list with 'Edinburgh' selected. Below the list is a 'Selection' field containing 'Edinburgh', a 'Prefix/search' field with 'Edinburgh', and an 'Update' button. The dialog also shows 'Web search' results for '110 hits' and a 'starting at' field. The right side of the dialog displays the selected term '2650225' and its definition: 'United Kingdom, Scotland, City of Edinburgh (seat of a first-order administrative division), lat/long 55.95206/-3.19648'. The dialog has buttons for 'Select term', 'Create term', 'Cancel', and 'Help'.

Classification clockmaker System CNAME
Classification System SHIC
BriefDescription
Type

Identification
Description
Aspect
Inscription
Production
Person
Place
PlaceIdentity
PlaceName
PlaceName
PlaceName
PlaceName
PlaceName
Date
Association
Period
Acquisition
Method
Person
Date
Disposal
Method
Date

Select from termlist

Terms

- Arrecife Edinburgh
- Barcelo Edinburgh Carlton Hotel
- Best Western Edinburgh City Hotel
- Cape Edinburgh
- Castle Edinburgh
- Channings Hotel Edinburgh
- City of Edinburgh
- Edinburg
- Edinburgh**
- Edinburgh (historical)
- Edinburgh [10]
- Edinburgh [11]

Selection

Prefix/search

Show non preferred terms

Web search

110 hits

Show starting at

HTML view Edit view

Term: 2650225

2650225 is the preferred term for **Edinburgh**.

Definition:
United Kingdom, Scotland, City of Edinburgh (seat of a first-order administrative division), lat/long 55.95206/-3.19648

SPARQL end-point as web termlist

The screenshot displays a web interface for a SPARQL end-point, showing a tree structure on the left and a detailed view of a selected term on the right.

STRUCTURE

- Date ▾
 - DateBegin ▾
 - DateEnd ▾
 - Note ▾
- Note ▾
- Field Collection**
 - Place
 - PlaceName ▾
 - PlaceName ▾
 - PlaceName ▾
 - PlaceName ▾
 - PlaceName ▾
 - SiteName ▾
- Date ▾
- Association**
 - Period ▾
 - Period ▾
 - Period ▾
 - Period ▾
 - Period ▾
 - Period ▾
- Acquisition**
 - Method ▾
 - ReferenceNumber ▾
 - Institution ▾
 - Number ▾
- Description**
 - Material ▾
 - Keyword ▾

DATA

Select from termlist

Terms

<http://collection.britishmuseum.org/id/thesauri/x11030>

Selection

Prefix/search

Show non preferred terms

Web search

Show starting at

HTML view

Concept:
<http://collection.britishmuseum.org/id/thesauri/x11030>

Preferred term: ivory

Definition:
IVORY and its Narrow Terms generally refer to worked material. See also TUSK and its Narrow Terms which generally refer to unworked material. The two terms have been used interchangeably in some records.

Broader concept:
<http://collection.britishmuseum.org/id/thesauri/x11187>

Related concept:
<http://collection.britishmuseum.org/id/thesauri/x10295>

Related concept:
<http://collection.britishmuseum.org/id/thesauri/x11462>

Related concept:
<http://collection.britishmuseum.org/id/thesauri/x11895>

<http://collection.britishmuseum.org/id/thesauri/x11030>

Control record for web termlist

The screenshot shows the 'Modes Admin' application window with the title 'Modes Admin: application Termlist selected'. The interface includes a menu bar (Controls, Application, File, Grid, Help) and a sub-menu bar (Application, File, Indexes, Validation, Grid, Views, Form entry, Report, Media, Termlist, Publication). A 'Termlists' dropdown menu is set to 'BM web termlist', with buttons for 'New Termlist desc.', 'Cancel changes', and 'Save changes'.

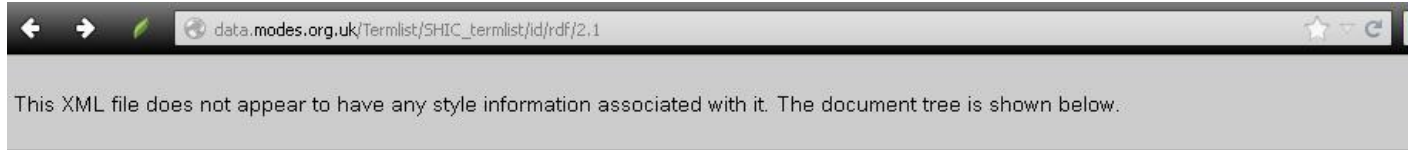
The main area displays a control record with the following fields:

control	
accessmethod	web
application	Termlist
+ file	bmmaterialscache
index	Terms
uindex	Non-preferred terms
candidateindex	Candidate terms
browseview	termlist\termsummary.xml
modesserver	
serverport	16000
username	
password	
url	http://collection.britishmuseum.org/id/thesauri/
webidentifiertomodes	termlist/rdf2termlist.xml
searchresulttomodes	termlist/rdf2termlist.xml
resultcountpath	
startparameter	
ufpath	ancestor::Material/Keyword[1]/text()
ufurl	http://collection.britishmuseum.org/Sparql?Query=PREFIX+skos%3A+%3Chttp%3A%2F%2F
accept	application/rdf+xml
status	Local control creating using Modes Complete v1.2

Modes Linked Data Framework

- Generic framework for publishing *any* Modes resource as Linked Data
- Based on a 404 Not Found handler (js)
- Control records state what can be published
- Uses XSLT to transform and deliver content
- Uses Linked Data API-like approach to URL structure
- Simple search; no SPARQL support

Standard Modes termlists



```
- <rdf:RDF>
- <skos:ConceptScheme rdf:about="http://data.modes.org.uk/Termlist/SHIC_termlist">
  <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#ConceptScheme"/>
  </skos:ConceptScheme>
- <skos:Concept rdf:about="http://data.modes.org.uk/Termlist/SHIC_termlist/id/2.1">
  <skos:inScheme rdf:resource="http://data.modes.org.uk/Termlist/SHIC_termlist"/>
  <skos:prefLabel xml:lang="en">2.1</skos:prefLabel>
  <skos:scopeNote xml:lang="en">DOMESTIC AND FAMILY ADMINISTRATION AND RECORDS</skos:scopeNote>
- <skos:scopeNote xml:lang="en">
  Includes all aspects of running the household, legal documents, accounts, bills, etc. and family records.
  </skos:scopeNote>
- <skos:scopeNote xml:lang="en">
  Example: e.g. title deeds, rent book, rate demand, eviction notice, account book, photograph album, loose family photographs, pedigree.
  </skos:scopeNote>
  <skos:broader rdf:resource="http://data.modes.org.uk/Termlist/SHIC_termlist/id/2"/>
  <skos:altLabel xml:lang="en">accounts</skos:altLabel>
  <skos:altLabel xml:lang="en">administration</skos:altLabel>
  <skos:altLabel xml:lang="en">documentation</skos:altLabel>
  <skos:altLabel xml:lang="en">family photography</skos:altLabel>
  <skos:altLabel xml:lang="en">family records</skos:altLabel>
  <skos:altLabel xml:lang="en">finance</skos:altLabel>
  <skos:altLabel xml:lang="en">genealogy</skos:altLabel>
  <skos:altLabel xml:lang="en">photographic records</skos:altLabel>
  <skos:altLabel xml:lang="en">records</skos:altLabel>
  </skos:Concept>
</rdf:RDF>
```

Shakespeare's plays

```
- <rdf:RDF>
- <rdf:Description rdf:about="http://richardlight.org.uk/Plays/shakespeare/id/636207">
  <rdf:type rdf:resource="http://richardlight.org.uk/play-schema.rdf#Speech"/>
  <play:inPlay rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet"/>
  <play:inScene rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.1.5"/>
  <play:isDeliveredBy rdf:resource="http://richardlight.org.uk/Person/shakespeare-characters/id/FathersGhost-ham"/>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.745">I am thy father's spirit,</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.746">Doom'd for a certain term to walk the night,</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.747">And for the day confin'd to fast in fires,</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.748">Till the foul crimes done in my days of nature</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.749">Are burnt and purg'd away. But that I am forbid</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.750">To tell the secrets of my prison house,</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.751">I could a tale unfold whose lightest word</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.752">Would harrow up thy soul, freeze thy young blood,</play:speechHasLine>
- <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.753">
  Make thy two eyes, like stars, start from their spheres,
</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.754">Thy knotted and combined locks to part,</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.755">And each particular hair to stand on end</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.756">Like quills upon the fretful porcupine.</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.757">But this eternal blazon must not be</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.758">To ears of flesh and blood. List, list, O, list!</play:speechHasLine>
  <play:speechHasLine rdf:resource="http://richardlight.org.uk/Plays/shakespeare/id/hamlet.line.759">If thou didst ever thy dear father love-</play:speechHasLine>
</rdf:Description>
</rdf:RDF>
```

Attack problem from both ends

- Publish as Linked Data
 - Shared ontologies, vocabularies, classifications
 - Resources (e.g. Shakespeare; museum catalogues)
- Consume Linked Data
 - Use “web termlists” to grab URLs for use alongside (or instead of) string-value data
- Source-agnostic: there is nothing Modes-specific about interface or data
- Adlib, CALM also developing web termlists

Dog food! Modes LD as web termlist

The screenshot displays a museum database interface. The background shows a record for a pocket watch with the following details:

- ObjectName:** simple name watch, other name pocket watch
- Classification:** decorative art, trade, 3.42
- Description:** Pocket watch, silver case
- Production:** clockmaker Dickie, Andrew; Place: Edinburgh, Lothian, Scotland; Date: 1866
- Acquisition:** Method: purchase, c; Date: 16.4.1930

The foreground shows a modal window titled "Select from termlist" with the following content:

- Terms:** A list containing "commission" and "copy".
- Selection:** A text input field containing "commission".
- Prefix/search:** A text input field containing "c", a checked "Show non preferred terms" checkbox, and an "Update" button.
- Web search:** "Search complete" and a "Web search" button.
- Term: commission**
 - Definition:** an original item made at the instigation of the museum
 - Related term:** [gift](#)
 - Related term:** [purchase](#)
- Buttons:** Select term, Create term, Cancel, Help