Connecting Data Objects using Linked Open Data

Lec Maj & Emmanuelle Delmas-Glass
Yale Center for British Art
What is YCBA-CIA working on?

• Began with access to **online collections** via website
• Extended YCBA **XML data provider**, data used by:
  - Discover Yale Digital Content
  - Google
  - ARTSTOR
  - Yale
  - YOUR PAINTINGS
  - [edition]
  - ART SY
  - artfinder

• Now focusing on building out semantic web: **connecting data** in more meaningful ways using linked data approach

"Connecting Data Objects" by Maj, Delmas-Glass @ YCBA Curatorial Discussion, 2012
What is Semantic Web / Linked Data?

• Sematic Web is a web of data objects, unlike current web of documents (publications about the objects).
• A “data object” can be tangible (as a painting) or concept representation (as unknown artist).
• “Data objects” are represented by identifiers called URIs (Uniform Resource Identifiers), not to be confused w/URLs (Uniform Resource Locator).
• We now connect manually different data objects to each other. Eventually automatically, and then via inference. Adding meaning to the relationship.

"Connecting Data Objects” by Maj, Delmas-Glass @ YCBA Curatorial Discussion, 2012
What is going on with this data?
Answer requires domain expert

- Image from Tate
- Data from YCBA
- Different objects
- BM, V&A, etc. have related objects as well

“Linked Data for researchers, YCBA Pilot Project” by Lec Maj @ LDBC2012
Better Data for Research

• Answer in-depth questions
• Refine and collect better data
• Current research
  – Spreadsheets
  – Paper print outs
  – Google search
  – Image search
• Research Questions
  – What related types of objects are in existence
  – What drawings were created and owned by Paul Sandby (analysis of image markings to identify marks on them)
  – History map of ownership migration for Sandby’s drawings
  – Compare compositions of water colors to other objects to help speculate artistic models (how was the work influenced, how it is treated by various artists)
• Tools
  – Need visual tools for domain experts to analyze data
  – Combination of maps, timelines, etc. without technical expertise and complicated tools

“Linked Data for researchers, YCBA Pilot Project” by Lec Maj @ LDBC2012
Example of Mrs. Abington as Miss Prue

YCBA Data Object representing painting

Painting
*Mrs. Abington as Miss P...*
Sir Joshua Reynolds, 172...
Oil on canvas

“Creation Event”

Google Data Object representing creator

Joshua Reynolds
Sir Joshua Reynolds RA FRSA was an influential 18th-century English painter, specialising in portraits and promoting the “Grand Style” in painting which depended on idealization of the imperfect.

Wikipedia

Born: July 16, 1723, Plymouth
Died: February 23, 1792, Richmond upon Thames
Parents: Samuel Reynolds
Education: Plympton Grammar School
Artwork: The Age of Innocence, Lady Cockburn and Her Three Eldest Sons, More
Siblings: Elizabeth Reynolds, Frances Reynolds, Mary Reynolds

“Exhibition Publication Event”

OCLC Data Object representing publication

“Connecting Data Objects” by Maj, Delmas-Glass @ YCBA Curatorial Discussion, 2012
How is Linked Data Formatted?

- Linked Data is represented in a file using any number of syntaxes: RDF, TTL, RDFa, etc.
- It contains RDF (resource description format) triples (statements):
  - Sky (subject) is (predicate) blue (object)
  - Painting (ObjectID: 499) is documented in Book (BIBID: 7838392)
  - Painting (ObjectID: 499) is documented in Book (OCLC #: 74916141)
Challenges we encountered

• Limitations with Cross Collection Discovery (CCD)
  – Flat index, no way to link multiple concepts, fast search

• Limitations in data entry
  – Can update only “generic” object
  – Cannot update specific part of a record, item, components
  – page level cataloging

• Limitation on how data can be linked

• Much of the work is manual process
  – Need to utilize Automatic extraction of Linked Data
  – Need more texts, OCR, run via 3rd party tools

http://collection.britishart.yale.edu/id/object/236

"Connecting Data Objects” by Maj, Delmas-Glass @ YCBA Curatorial Discussion, 2012
More complicated example: *In and out of Love*

<table>
<thead>
<tr>
<th>Creator</th>
<th>Damien Hirst, born 1965, British;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td><em>In and Out of Love (Butterfly Paintings and Ashtrays)</em></td>
</tr>
<tr>
<td>Date</td>
<td>1991</td>
</tr>
<tr>
<td>Medium</td>
<td>Eight paintings, household paint on canvas with butterflies; four boxes; one table, formica top on steel base; four glass ashtrays filled with cigarette butts</td>
</tr>
<tr>
<td>Dimensions</td>
<td>(eight paintings): 60 x 60 inches (152.4 x 152.4 cm) Overall (four boxes): 36 x 36 x 36 inches (91.4 x 91.4 x 91.4 cm) Overall (one table): 42 inches (106.7 cm)</td>
</tr>
</tbody>
</table>

URI: [http://collection.britishart.yale.edu/id/object/4908](http://collection.britishart.yale.edu/id/object/4908)

In 1991, for his first solo exhibition, Damien Hirst transformed a disused two-story travel agent on Woodstock Street in east London into his landmark...

"Connecting Data Objects" by Maj, Delmas-Glass @ YCBA Curatorial Discussion, 2012
Linked data graph for Damien Hirst’s installation *In and Out of Love*

"Connecting Data Objects" by Maj, Delmas-Glass @ YCBA Curatorial Discussion, 2012
Linked Data (LD) Benefits

- RDF (resource description format) – attaches meaning to data and content
- Does not matter what kind of data it is: structured, unstructured, or in between
- RDF forges semantic link between smallest components of data so they can become part of an answer to query
- break down the walls and links ALL data together
- Merge private and public data
- Enables scholars, students, researchers, companies and people in general to do more, generate innovation from data (answer more questions, give more meaningful answers, useful applications, tools, etc.)

"Connecting Data Objects" by Maj, Delmas-Glass @ YCBA Curatorial Discussion, 2012
What is the data / research that YCBA wants to be authority over and publish?

- Art Objects (Paintings, Frames, etc.)
- Rare Books
- Library & Archive
- Exhibitions
- YCBA produced publications
- Louis Kahn building data
- Photo Archive

"Connecting Data Objects" by Maj, Delmas-Glass @ YCBA Curatorial Discussion, 2012
How do we get it done?

1. Generate and publish YCBA material
   – part of online collections
   – web pages via Drupal Content Management System
   – PDF files that can be transformed into linked data

2. Assign identifiers for data objects
   – URIs
   – verify data objects

3. Link data objects
   – manually (for verification of relationship data types)
   – Automatically (extracting data from exhibition catalogs)
   – inference

"Connecting Data Objects" by Maj, Delmas-Glass @ YCBA Curatorial Discussion, 2012
Automatic Linking using OpenCalais

http://viewer.opencalais.com/

"Connecting Data Objects" by Maj, Delmas-Glass @ YCBA Curatorial Discussion, 2012
Connecting Data Globally

CIDOC-CRM Ontology
V&A
BM
YCBA
BL
YCBA Lib

As of September 2011

Is it all data files and acronyms?
Visual SPARQL Tools

“Cross organization data harmonization and access, linked data decentralized approach” Maj, Mahmud, Delmas-Glass – CIDOC 2012
Why are we engaged in LOD?

Rembrandt Federated Search Results (Sofia - OWLIM)

- Rembrandt Database: Saskia van Uylenburgh in arcadian costume
  http://www.rembbrandtdatabase.org/obj/22578
- Rembrandt Database: Man in Oriental Costume
  http://www.rembbrandtdatabase.org/obj/50050

Yale Federated Search Results (New Haven - OWLIM)

- Yale Database: Costume Design
  http://collection.britishart.yale.edu/id/object/10031
- Yale Database: The Countess of Goring in the Costume of the Chorlton Hunt
  http://collection.britishart.yale.edu/id/object/1004
- Yale Database: Costume Study for a Portrait of a Man
  http://collection.britishart.yale.edu/id/object/12592
- Yale Database: Costume Study for the Portrait of a Lady
  http://collection.britishart.yale.edu/id/object/14141
- Yale Database: Study of a Gentleman in 17th Century Costume
  http://collection.britishart.yale.edu/id/object/15123
- Yale Database: Figures in Totemic Costume and Totem Pole, Number 149
  http://collection.britishart.yale.edu/id/object/20971

Courtesy: Dominic Oldman @ British Museum, Research Space
Maps based on LOD
Related Art and Book objects