RDF Data and Image Annotations in ResearchSpace

By Jana Parvanova, Vladimir Alexiev, Stanislav Kostadinov

DH-CASE 2013
Florence, Italy
The ResearchSpace Project

- Aims to provide a collaboration environment for research projects in the humanities area
- The British Museum manages the project
- Ontotext implements the application and provides semantic technology expertise; OWLIM is used as a semantic repository
- Proposed and Funded by the Andrew W. Mellon Foundation
- Part of a programme of Mellon Foundation Projects: CollectionSpace, ConservationSpace, ResearchSpace
- Stage 3 (working prototype) was developed in 2011-2012
- Stage 4 is expected to start in 2013
Data and Tools

Conversion

Existing data sets (collections, open linked data etc, images)

CIDOC-CRM data

ResearchSpace

New data records

Annotations, links and other user-generated content

Publication

RS Tools: semantic search, data annotation, image annotation, data basket, approval workflow
Data Annotation

As1979.01.53 (BM Registration Number)
(Nuxeo UID : BM-RRM186876)

Current owner
The British Museum

Current keeper
The British Museum

BM Department Middle East

Former or current owner
University of Cambridge

Former or current keeper
University of Cambridge

Object type
dress

Material
silk
silver

Production & Techniques
used general technique embroidered
took place at Aleppo (Made)

Dimensions
Length : 115.00 cm
Width : 130.00 cm (across arms)

Annotations

New Annotation

Title: dress RRM186876 - silk
Referred value:
silk
Suggested value:

After the fall of the Byzantine empire, the silk industry began to decline in Syria. The industry saw a huge international expansion, which caused the demand and the prices of silk to decrease drastically, no longer making it a profitable venture. This decline was again furthered by the introduction of the mechanical loom into the textile industry.

Add topic...
Data Annotation

As1979.01.53 (BM Registration Number)
BM-RRM186876

Current owner
The British Museum

Current keeper
The British Museum
BM Department Middle East

Former or current owner
University of Cambridge

Former or current keeper
University of Cambridge

Object type
dress

Material
silk
silver

Production & Techniques
used general technique embroidered
took place at Aleppo (Made)

Dimensions
Length: 116.00 cm
Width: 130.00 cm (across arms)

Annotations

dress RRM186876 - silk
Referred value: silk
2013-09-12 11:54 AM, Administrator
After the fall of the Byzantine empire, the silk industry began to decline in Syria. The industry saw a huge international expansion, which caused the demand and the prices of silk to decrease drastically, no longer making it a profitable venture. This decline was again furthered by the introduction of the mechanical loom into the textile industry.

silk production
Syria, North East

2013-09-12 11:57 AM, Administrator
Same traditional patterns may be observed here: RRM186883
woman's scarf (embroidered, with Arabic inscription)

Object: RRM186883 Woman's scarf (embroidered, with Arabic inscription)
Hot water dyes, including both acid dyes and hot water fibre reactive dyes, can, instead of steaming, be applied by immersing the fibre to be dyed in a hot dye bath, usually for half an hour or an hour, along with any required auxiliary chemicals.
Image Annotation

Annotations

- AN00343 - embroidery
  2013-09-12 12:08 PM, Administrator
  Traditional embroidery was typically produced by village and Bedouin women rather than town dwellers, who have usually worn Western or Ottoman attire. This fact probably explains why most of the research done to date on costumes and embroidered art has focused mainly on the rituals and social ceremonies in rural and Bedouin areas.

- AN00343 - high quality silk threads
  2013-09-12 12:09 PM, Administrator
  Hot water dyes, including both acid dyes and hot water fibre reactive dyes, can, instead of steaming, be applied by immersing the fibre to be dyed in a hot dye bath usually for half an hour or an hour, along with any required auxiliary chemicals.
Implementation: Data models

- **CIDOC-CRM**: provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation.

- **OAC**: An Annotation is considered to be a set of connected resources, typically including a body and target, where the body is somehow about the target.

- **SKOS**
Implementation: Example

CIDOC-CRM

OAC
Implementation: Example
Implementation - Images

Posted on: 04/06/2013
By: M. Schwarz
The lower ground colour of his canvasses was often orange-red earth, the upper ground layer of ochre....

Re: Colors
Rembrandt did not as a rule do any preliminary sketches...

Shapes are annotated and annotations are saved in OWLIM triple store using Open Annotation data model.

Deep zoom server serves tiles to a JavaScript image viewer.

Users draw shapes to point to specific parts of the image.
Some Numbers

- Museum objects: 2,051,797
- Thesaurus entries: 415,509 (a total of 90 ConceptSchemes)
- Explicit statements: 195,208,156. We estimate that of these, 185M are for objects (90 statements per object) and 9M are for thesaurus entries (22 statements per term).
- Total statements: 916,735,486. The expansion ratio is 4.7x (i.e. for each statement, 3.7 more are inferred). This is considerably higher compared to the typical expansion for general datasets (e.g. DBpedia, GeoNames, FactForge) that is 1.2 - 2x.
- Nodes: 53,803,189. Includes unique URLs and literals (this dataset doesn’t use blank nodes)
- Repository size: 42 Gb, object full-text index: 2.5 Gb, thesaurus full-text index (used for search auto-complete): 22Mb.
Wrap-up

• Future plans
• Questions?
• Contacts: jana.parvanova@ontotext.com, vladimir.alexiev@ontotext.com
• Links:
  • http://www.researchspace.org/
  • The British Museum, CIDOC CRM and the Shaping of Knowledge - by Dominic Oldman, principal investigator
  • Implementing CIDOC CRM Search Based on Fundamental Relations and OWLIM Rules – by Vladimir Alexiev
  • ResearchSpace CIDOC CRM Search System – screencast on YouTube
• Thank you!