Museum data, where next – consuming linked data

Cultural Heritage and the Semantic Web
British Museum, UK, 13th January 2011

Hugh Glaser
Seme4 Limited and University of Southampton, UK
From the Study Day Description

• “The ability of the semantic web to cheaply but effectively integrate data and breakdown data silos provides museums with a long awaited opportunity to present a richer, more informative and interesting picture.”

  » Dominic Oldman
### Ordnance Survey Linked Data

A description of the resource identified by [http://data.ordnancesurvey.co.uk/](http://data.ordnancesurvey.co.uk/)

<table>
<thead>
<tr>
<th>Title</th>
<th>Ordnance Survey Linked Data.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Ordnance Survey is Great Britain's national mapping agency, providing the most accurate and up-to-date geographic data, relied on by government, business and individuals. OS OpenData is the opening up of Ordnance Survey data as part of the drive to increase innovation and support the &quot;Making Public Data Public&quot; initiative. As part of this initiative Ordnance Survey has published a number of its products as Linked Data. Linked Data is a growing part of the Web where data is published on the Web and then linked to other published data in much the same way that web pages are interlinked using hypertext. The term Linked Data is used to describe a method of exposing, sharing, and connecting data via URIs on the Web. To find more Linked Data published as part of this initiative please go to data.gov.uk. If you are not familiar with Linked Data, OS OpenData products are also available in alternative formats from the OS OpenData website. Ordnance Survey can provide support for the Ordnance Survey OpenData products, but cannot give advice or support on using RDF, SPARQL or SPARQL Endpoints. Ordnance Survey has published three separate linked data resources: the 1:50 000 Scale Gazetteer, Code-Point Open and the administrative geography gazetteer for Great Britain.</td>
</tr>
<tr>
<td>Creator</td>
<td>Ordnance Survey</td>
</tr>
<tr>
<td>Publisher</td>
<td>Ordnance Survey</td>
</tr>
<tr>
<td>Date issued</td>
<td>2010-10-25</td>
</tr>
<tr>
<td>Example resource</td>
<td>S016.4GU</td>
</tr>
<tr>
<td></td>
<td>Portswood</td>
</tr>
<tr>
<td></td>
<td>Southampton</td>
</tr>
<tr>
<td></td>
<td>Southampton, Abstract Geometry</td>
</tr>
<tr>
<td></td>
<td>Southampton, Itchen</td>
</tr>
<tr>
<td></td>
<td>The County of Hampshire</td>
</tr>
<tr>
<td></td>
<td>The City of Southampton</td>
</tr>
<tr>
<td></td>
<td>Southampton Common</td>
</tr>
<tr>
<td>SPARQL endpoint</td>
<td>Sparql</td>
</tr>
<tr>
<td>UIRT lookup point</td>
<td>Data/about=</td>
</tr>
<tr>
<td>Vocabulary used</td>
<td>Spatial Relations Ontology</td>
</tr>
<tr>
<td></td>
<td>The administrative geography and civil voting area ontology</td>
</tr>
<tr>
<td></td>
<td>Geometry Ontology</td>
</tr>
<tr>
<td></td>
<td>FIDAP</td>
</tr>
<tr>
<td></td>
<td>OWL</td>
</tr>
<tr>
<td></td>
<td>Gazetteer Ontology</td>
</tr>
<tr>
<td></td>
<td>Postcode Ontology</td>
</tr>
<tr>
<td>UIRT regex pattern</td>
<td><a href="http://data.ordnancesurvey.co.uk/ld/+">http://data.ordnancesurvey.co.uk/ld/+</a></td>
</tr>
<tr>
<td>See also</td>
<td>Semantic.html</td>
</tr>
<tr>
<td>Coverage</td>
<td>England</td>
</tr>
<tr>
<td></td>
<td>Scotland</td>
</tr>
<tr>
<td></td>
<td>Wales</td>
</tr>
</tbody>
</table>
BBC World Cup 2010 dynamic semantic publishing

Jem Rayfield | 10:00 UK time, Monday, 12 July 2010

The World Cup 2010 website is a significant step change in the way that content is published. From first using the site, the most striking changes are the horizontal navigation and the larger, format high-quality video. As you navigate through the site it becomes apparent that this is a far deeper and richer use of content than can be achieved through traditional CMS-driven publishing solutions.

The site features 700-plus team, group and player pages, which are powered by a high-performance dynamic semantic publishing framework. This framework facilitates the publication of automated metadata-driven web pages that are light-touch, requiring minimal journalistic management, as they automatically aggregate and render links to relevant stories.
The Rosetta Stone

Object type
stein (all objects)

Title (object)
The Rosetta Stone

Materials
granodiorite (all objects)

Place (findspot)
Excavated/Findspot: Fort Saint Julien (all objects)

Date
196BC

Period/Culture

Provenance: (scope note | all objects)

Description
Part of grey and pink granodiorite stele bearing priestly decree concerning Ptolemy V in three blocks of text: Hieroglyphic (14 lines), Demotic (32 lines) and Greek (53 lines).

Inscriptions
Inscription Type: inscription
Inscription Language: Greek

Dimensions
Length: 112.3 centimetres (max)
Width: 78.7 centimetres
 Thickness: 28.4 centimetres

Condition
Fair (incomplete)

Curator’s comments
Compass text:
The Rosetta Stone

Object types:
- stele (all objects)

Title (object):
The Rosetta Stone

Materials:
granodiorite (all objects)

Place (findspot):
Excavated/Found at Fort St. Julien (all objects)

Date:
196BC

Period/Culture:
Prolemaic (scope note | all objects)

Description:
Part of grey and pink granodiorite stele bearing priestly decree concerning Ptolemy V in three blocks of text: Hieroglyphic (14 lines), Demotic (32 lines) and Greek (53 lines).

Inscriptions:
- Inscription Type: inscription
- Inscription Language: Greek

Dimensions:
- Length: 112.3 centimetres (max)
- Width: 78.7 centimetres
- Thickness: 28.4 centimetres

Department: Ancient Egypt & Sudan
Registration number: 24

Seme4.com

© The Trustees of the British Museum
Description
Part of grey and pink granodiorite stele bearing priestly decree concerning Pharaoh V in three blocks of text: Hieroglyphic (14 lines), Demotic (12 lines) and Greek (53 lines).

Inscriptions
Inscription Type: Inscription
Inscriptions Language: Greek

Dimensions
Length: 112.3 centimetres (max)
Width: 75.7 centimetres
Thickness: 28.4 centimetres

June 1999 – Temporary Exhibition

Treatment date
24 June 1999

Reason for treatment
Temporary Exhibition

Treatment proposal
1.TextNode cleaned return edges. 2. Clean and prepare for “Rosetta Stone, Cracking Codes and Decipherment” exhibition.

Condition
Object coated with protective treatment - identified as camouba wax, hand-grease and dirt. Inscriptions bear traces of hard pink material, particularly near edges and at corners. It closely matches the pink granitic vein for colour and is covered in places by black patina ink – see analysts report. Inscription was painted with water-based white Phala gouache, by Carol Andrews in 1982 to replace white powder. (White powder was probably modelling clay applied as a slip over entire surface; this excess would have been removed with a cloth). While powder (clay) continued to absorb areas of damage, returns at either side have while deposit with pronounced horizontal directionalcy, but areas of damage do not have this deposit. (This is possibly gypsum plaster, spilled during casting of the inscription in Egypt). Also found: traces of reddish compacted clay on base and returns and a looser whitish material under cross-member of lid mount, but elsewhere consolidated and turned brown by camouba wax. Text painted in white on left return, “CAPTURED IN EGYPT BY THE BRITISH ARMY IN 1801”, on right return, “PRESENTED BY KING GEORGE III”.

Treatment details
Samples were taken of pink material and printers ink from inscription; surface coating from various locations, white deposit on returns, reddish compacted clay on base and returns: loose whitish
British Museum

• That was a year ago
  – Science data
  – Conservation data
  – Linked to the Catalogue
  – Live on the web site

• Now
  – Catalogue

• ResearchSpace
  – More institutions
  – Research environment
The LOD “cloud”

Our Example: RKB and RKBExplorer

- RKB (ReSIST Knowledge Base) and RKBExplorer
- Knowledge-enabled infrastructure for cooperation in research into resilient systems
- Reasonably mature system
RKBExplorer – the ReSIST Project

ReSIST (Resilience for Survivability in IST) is a NeE that addresses the strategic objective "Towards a global dependability and security framework" of the Work Programme, and responds to the stated "need for resilience, self-healing, dynamic content and volatile environments". It will integrate leading researchers active in the multidisciplinary domains of Dependability, Security, and Human Factors, in order that Europe will have a well-focused coherent set of research activities aimed at ensuring that future "ubiquitous computing systems", the immense systems of ever-evolving networks of computing and mobile devices which are needed to support and provide Ambient Intelligence (AI), have the necessary resilience and survivability, despite any residual development and physical faults, interaction mistakes, or malicious attacks and disruptions. The objectives of the Network are: 1) integration of teams of researchers so that the fundamental basis
But Why?

www.rkbexplorer.com
<sameAs>
interlinking the Web of Data

The Web of Data has many equivalent URIs. This service helps you to find co-references between different data sets. Enter a known URI, or use Sindice to search first.

Why not try searching for the string “Southampton”, (which we will look up for you on Sindice first) or finding other equivalent identifiers for http://transport.data.gov.uk/id/local-authority/1755?

Currently serving 39,221,073 URIs in 13,112,833 bundles!

about · contact · get the widget

© 2010 Seme4
Seme4.com
<sameAs> interlinking the Web of Data

The Web of Data has many equivalent URLs. This service helps you to find co-references between different data sets. Enter a known URI, or use Sindice to search first.

Equivalent URLs for http://acm.rkexplorer.com/id/person-344165-529a2387779b9822d610a72843a3980c –

2. http://acm.rkexplorer.com/id/person-344165-529a236776b9822d610a72843a3980c
3. http://acm.rkexplorer.com/id/person-344165-be4c940a0f0b483a2d1d75df9fa0bab6f
4. http://acm.rkexplorer.com/id/person-344165-c23ba3f792faeac763400f197b999ad1b
5. http://acm.rkexplorer.com/id/person-344165-e0b9eea0f283efab15e86ae81bf111
7. http://dblp.rkexplorer.com/people-2ceba813b69a235014d3d46c881a1...88ce3c9dc7
8. http://dblp.rkexplorer.com/people-cd93ebab4b98a73185165ff5f...f1785e51d3
9. http://dblp.rkexplorer.com/people-cd93ebab4b98a73185165ff5f...88cd85ab7e
10. http://dblp.rkexplorer.com/people-cd93ebab4b98a73185165ff5f...601d0b236
11. http://dblp.rkexplorer.com/people-cd93ebab4b98a73185165ff5f...93f9b9e9f1
12. http://dblp.rkexplorer.com/people-cd93ebab4b98a73185165ff5f...872498927e
13. http://dblp.rkexplorer.com/people-cd93ebab4b98a73185165ff5f...411b54ff75
14. http://acm.rkexplorer.com/id/resource-CSP162937-4be60e02ed76f...904295f72
17. http://acm.rkexplorer.com/id/resource-CSP162937-83b9db880b
18. http://acm.rkexplorer.com/id/resource-CSP162937-8ac5adb3359f3...213751b546
What is Southampton?
Co-Reference

- Repositories have people, publications, etc. from other institutions who also have records there and elsewhere

- And vice versa

**Co-Reference is a Big Problem**
- Everything is a URI (not title, name, number…)
- Identifying multiple URIs for one resource
- Rejecting incorrectly conflated resources
- Publishing
- Using

**Coldstart**
- A serious problem
- Nothing is linked to anything
- Not even (reliably) within most repositories
Co-Reference Service (CRS)

- **CRS Subsystem**
  - Find co-references
  - Store them
  - Publish them
    - Essentially:
      - $\text{URI}_i \rightarrow \{ \text{URI}_1, \ldots, \text{URI}_i, \ldots, \text{URI}_n \}$
    - Recommend a “Canon”
- **Published by the Data Publisher**
  - And possibly others
- **Middleware aggregates co-references from recognised CRses**
CRS continued

- **CRS Policies are defined by context**
  - Often one per Triplestore
  - Can be many per Triplestore for different purposes
  - May not be associated with a particular Triplestore

- **Maintenance**
  - Provenance
  - Rollback

- **Can be used to infer owl:sameAs**

- **Eg OAI CRS has**
  - 7531045 different URIs
  - in
  - 2544955 bundles
Open System

• **RKBExplorer is only one interface**
  – And not a required part

• **Services:**
  – Details for a paper (the right hand pane in RKBExplorer):
  – Network of people for a publication (lower pane):
  – ...

• **Services**

• **Other Interfaces (using the services)**
  – Personal Web pages
  – iPhone
  – iGoogle Gadget
This page gives convenient access to various useful links and demonstrations associated with RKBExplorer. It is primarily for the use of project members and collaborators, and not intended as a general entry point for other people. It exposes many of the services of the RKB system in RESTful-style manners. One day I will find the time to make a better page, with documentation.

If you are interested in something here and want to explore using it, please contact Hugh Glaser or Ian Millard.

RKB Explorer application
- ReSiST project, Jean-Claude Laprie, ALRL Paper

iGoogle Gadgets of the RKB Panels
- [Main Page]

Complete Co-Reference Information
- [Main Page], or try: Nigel Shadbolt, Hugh Glaser, Brian Randell, Jean-Claude Laprie, Seungwoo Lee

Ontology Mapping
- [Main Page], or try: KISTI -> AKT (RDF, html)

Networks (Communities of Practice)
- [Find the Network of a URI], or try: Hugh Glaser, The Semantic Web Revisited, Tim Berners-Lee

Detail Information
- [Find Information about a URI], or try: Hugh Glaser, The Semantic Web Revisited, Tim Berners-Lee

Why? - Why are two Things Related?
- [Main Page]
MyExperiment :: dotAC Explorer

Location of Project Members

Details

Names:
MyExperiment

Starts:
01/03/2007

Ends:
31/03/2009

Web Address:
http://www.myexperiment.org/

Description:
The myExperiment Virtual Research Environment enables you and your research

Participating Organisations

- University of Southampton

Related JISC Projects

- Meeting Memory Technology, Informing Collaboration
- myExperiment Repository Enhancement
- Collaborative Orthopaedic Research Environment
- Internet-Based Behavioural Research
- Virtual Research Integration Collaboration
- Building Research and Innovation Networks

Project Members

- Carla Goble
- Darius Michaelides
- David De Roure
- Frederique VanTill
**iPhone App**

**Hugh Glaser**

- **Telephone**: +44-1703-593670
- **Works For**: School of Electronics and Computer Science, University of Southampton
- **Full Name**: Hugh Glaser
- **Fax**: +44-1703-593045
- **Web**: [http://www.ecs.soton.ac.uk/people/hg](http://www.ecs.soton.ac.uk/people/hg)

**Projects**
- ReSIST Resilience for Survivability in IST
- HELIOS

**Research Areas**
- Static Analysis
Semantic Web is not “All or Nothing”
“The ability of the semantic web to cheaply but effectively integrate data and breakdown data silos provides museums with a long awaited opportunity to present a richer, more informative and interesting picture.”

» Dominic Oldman
Concluding Remarks

• RKB: Nobody had to put their personal data in – they come and find it there

• The initial British Museum project has delivered value, at comparatively low cost
• A new project is almost complete – the catalogue is in Linked Data form

• Linked Data works pretty well – especially for integration

• Added value
  – More sophisticated services
  – In an Open System you don’t have to do everything yourself
  – Concentrate on core business

• Don’t stop at the catalogue – there is a lot of other interesting stuff

• Worry about your co-reference
• A Whirlwind tour of a part of it that has hopefully given you a flavour

• Imagine people having an RKB-like worldwide view of Cultural Heritage
  – With your institution as part
You are never alone

- Ian Millard
- ReSIST Project
- AKT Project
- EnAKTing Project
- 10 years of collaborators

http://eprints.ecs.soton.ac.uk/id/eprint/21894