Consuming Linked Data

Hugh Glaser
Seme4 Limited and University of Southampton

http://eprints.ecs.soton.ac.uk/id/eprint/22613
Down & Dirty - URIs

http://imgs.xkcd.com/comics/standards.png
Co-Reference

- **Lots of resources with multiple URIs**
  - This is the nature of the Beast
  - Legal, Sociological and Technical reasons

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

- **Solution**
  - Co-reference is just Knowledge, but quite special, and crucial, knowledge
  - Generating a new URI is not the solution to already having too many
  - The web is anarchic – don’t try to impose impossible structure on it
  - Embrace the anarchy and multiple URIs in your systems ab initio
Co-Reference Service (CRS)

• **CRS Subsystem**
  – Find co-references
  – Store them
  – Publish them
  • Essentially:
    • URI<sub>i</sub> -> { URI<sub>1</sub>, ..., URI<sub>i</sub>, ..., URI<sub>n</sub> }
  – 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
Look up Seungwoo

kisti.rkbexplorer.com

Canon
http://kisti.rkbexplorer.com/id/PER_00000000000000131417

Duplicate URIs
http://acm.rkbexplorer.com/id/person-407157-c23d3b44f74d68c8e281e3ff7795bf4c
http://citeseer.rkbexplorer.com/id/resource-CSP180445-4e6f59a7e2af5838836e83e27748b7ac
http://dblp.rkbexplorer.com/id/people-1ec5a600299222dd6374695ef5214f05-3aac82bb29cfebfc2e83d8b23349445
http://kisti.rkbexplorer.com/id/PER_00000000000000131417

Bundle instantiated at
2009-12-20 02:39:10

Alternative representations
RDF/XML
sameAs list in N3
image/png
Co-Reference Closure

Complete Co-Reference Information

This service computes the equivalence class within the known URIs for a specified URI, by consulting all relevant CRS knowledge bases.

Display all related triples ? ☑

Display connectivity image? ☑ None □ JPEG □ PNG □ GIF □ Postscript □ SVG

Equivalent URIs...

1. (Canon) http://acm.rkibexplorer.com/id/person-407157-c23d3b4df746d05c2b81e3f77595b4c
2. http://iclass.rkibexplorer.com/id/resource-CSI180445-4a8f15a0a2a65389fa55a2774bf87a
3. http://iclass.rkibexplorer.com/id/resource-CSI180445-e8a83f83781f85b82395ee00b655
4. http://dblp.rkibexplorer.com/id/people-1ec50a062d9223d87469e6f5214005-5c14713563977ed31471469e6bb41
5. http://dblp.rkibexplorer.com/id/people-74c98bcb6d125385a66683924264b43b3c9c2d9e185c36e91e36323

The following diagram shows the interconnectivity between the CRS knowledge bases which maintain the context-dependent representation of co-referent for each of the RKIBExplorer domains.
<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.rpcexplorer.com/id/person-344165-529a23877796b9822d610a72843a3980c –
2. http://acm.rpcexplorer.com/id/person-344165-529a23877796b9822d610a72843a3980c
3. http://acm.rpcexplorer.com/id/person-344165-be4c9d40af0b48b3a2d1675d9f0a0bba8f
4. http://acm.rpcexplorer.com/id/person-344165-c23ba3f792faeaca753400f197b69ad1b
5. http://acm.rpcexplorer.com/id/person-344165-e0d9e6a0f283ef4bc5e3e86ec681f111
7. http://dblp.rpcexplorer.com/people-2e68b138f69e235045d33d45dc8b1a1...8dce3c39e7
8. http://dblp.rpcexplorer.com/people-9093eb40b98f7e3185165f5f9...7185f1d3
9. http://dblp.rpcexplorer.com/people-9093eb40b8a9e73185165f5f9...88ed85ab7e
10. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
11. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
12. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
13. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
14. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
15. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
17. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
18. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
19. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
20. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
21. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
22. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
23. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
24. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
25. http://dblp.rpcexplorer.com/people-9093eb40b98f73185165f5f9...601db0236
<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.

<sameAs> http://data.archiveshub.ac.uk/id/person/gb248/smithadam1723-1790pol...leconomist -

1. http://dbpedia.org/resource/A_Smith
2. http://dbpedia.org/resource/Adam_Smith
7. http://dbpedia.org/resource/Adam_Smith#Quotes
10. http://mpii.de/yago/resource/Adam_Smith
12. http://rdf.freebase.com/ns/guid.9202a8c0400064118000000000000065ab
15. http://sw.opencyc.org/concept/Mx4rwSxnQZwpEbGdrcN5Y29ycA
17. http://sw.opencyc.org/2009/04/07/concept/Mx4rwSxnQZwpEbGdrcN5Y29ycA

rdf+xml · n3 · json · text · show fewer items
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.

This is a sub-service of the sameAs service, and contains only information that is provided by a third party.

The data presented here relates to Hugh’s experiment in gathering together library subject and associated data from good sources. It is likely to form part of a forthcoming project in the library area.

Equivalent URIs for http://d-nb.info/gnd/4126114-8 –

1. http://bnb.data.bl.uk/id/concept/csh/Nazis
3. http://id.loc.gov/authorities/sh85090140#concept
Where Was I? :- Consuming Linked Data

• **Linked Data Browsers?**
  – Just soooooo boring
  – And not real user

• **Semantic Web Applications?**
  – Revolution, not evolution
  – Confuse presentation and data

• **Think about other people consuming**
  – Publishing isn’t exciting any more

• **Provide appropriate services in appropriate formats**
  – Consumable by “normal” systems
  – This means
    • Useful – doing useful things
    • RESTful – easy to access
    • JSON – easy to process
    • …
Wendy Hall is connected to Les Carr

They are linked by 127 relations.

**Publications**

They have co-authored 120 papers:
- Enhancing the Distributed Link Service for Multimedia and Collaboration
- Citation Linking: Improving Access to Online Journals
- Web Journals Publishing: a UK Perspective

(117 more)

**Affiliations**

They are both affiliated to 2 organizations:
- School of Electronics and Computer Science
- http://kisti.rkbexplorer.com/id/INS_0000000000000006193

**Projects**

They are both members of 3 projects:
- QUERIES IN CONTEXT (QUIC)
- Visiting Fellowship - Professor Ted Nelson
- COHSE: INFORMED WWW LINK NAVIGATION USING ONTOLOGIES

**Projects**

They are both contacts for the Collaborative Orthopaedic Research Environment project.
In the OAI.rkbexplorer.com Repository
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**
  - [http://www.rkbexplorer.com/services/](http://www.rkbexplorer.com/services/)
Semantic Web is not "All or Nothing"
Professor David C De Roure

School of Electronics and Computer Science
University of Southampton
Southampton
SO17 1BJ
United Kingdom

Positions: Head of Group in Grid and Pervasive Computing Group, Academic staff in Intelligence, Agents, Multimedia Group

Extension: 22418
Telephone: +44 (0)23 8059 2418
Email: dder@ecs.soton.ac.uk
URI: http://id.ecs.soton.ac.uk/person/47 [browse]

Interests: agent-based computing, agent-oriented computing, art and humanities applications, audio, augmented computing, cheminformatics, collaborative tools, computers and music, decentralised systems, distributed programming languages, distributed systems, dr who, e-learning, e-research, e-science, e-social science, electronic lab notebooks, electronic systems design, environmental monitoring, folklore studies, grid, grid computing, grid standards, hardware-software co-design, hypertext information systems, intelligent sensor networks, interaction, iot, linux, lisp, middleware, multi user dungeons, multicast, multimedia, music, music information retrieval, nature inspired computing, open access, open hypermedia, open journals, open source, pervasive computing, provenance, rdf, recommender systems, scheme, scholarly communications, scripting languages, self-organising systems, semantic annotation, semantic grid, semantic web, sensor networks, service-oriented architectures, simulation, social networks, social tagging, surge, technology-enhanced learning, ubiquitous computing, virtual research environments, visual programming, web, web 3.0, web science, web services, web standards, workflow

The group secretary for the Intelligence, Agents, Multimedia Group is Jane Morgan.

Biography

David De Roure is a Professor of Computer Science in the School of Electronics and Computer Science at the University of Southampton, UK. A founding member of the School's Intelligence, Agents, Multimedia Group, he leads the e-Research activities and is a Director of the Pervasive Systems Centre. Closely involved in the UK e-Science programme, David is National Strategic Director of e-Social Science, Chair of OMII-UK and a Co-Director of e-Research South. His work focuses on creating new research methods in and between diverse disciplines, and his projects draw on Web 2.0, Semantic Web and workflow technologies.
David De Roure is connected to Danius Michaelides

They are linked by 57 relations.

**Publications**

They have co-authored 52 papers:

- Semantic Annotation in Ubiquitous Healthcare Skills-based Learning Environments
- A Card Based Metaphor for Organising Pervasive Educational Experiences
- Supporting Domain Experts in Creating Pervasive Experiences

(49 more)

**Affiliations**

They are both affiliated to 2 organizations:

- School of Electronics and Computer Science
- University of Southampton

**Projects**

They are both contacts for 2 projects:

- Meeting Memory Technology Informing Collaboration
- MyExperiment

**Organisation**

They are both affiliated to: University of Southampton.
The Rosetta Stone

Object types:
- stele (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 (52 lines) and Greek (53 lines).

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

- Inscription Type: inscription
- Inscription Script: Hieroglyphic
- Inscription Language: Greek

- Inscription Type: inscription
- Inscription Script: Demotic

- 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 type(s):
stele (all objects)

Title (object):
The Rosetta Stone

Materials:
granodiorite (all objects)

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

Date:
196BC

Period/Culture:

Provenance:

Description:
Part of grey and pink granodiorite stele bearing Priestly decree concerning Ptolemy V in three blocks of text: Hieroglyphic (14 lines), Demotic (12 lines) and Greek (53 lines).

Inscriptions:
Inscription Type: inscription
Inscription Language: Greek

Inscription Comment: The inscription is a decree passed by a council of priests, one of a series that affirm the royal cult of the 13-year-old Ptolemy V on the first anniversary of his coronation.

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

Registration number: 24

Additional ID:
B3.24

Bibliographic reference:

June 1999 – Temporary Exhibition

April 1992 – Examination
Description
Part of grey and pink granite stela, stela bearing priestly decree concerning Pharaoh V in these blocks of text: Hieroglyphic (14 lines), Demotic (12 lines) and Greek (53 lines).

Inscriptions
Inscriptions Type: inscription
Inscriptions Language: Greek

Inscriptions Comment: The inscription is a decree passed by a council of priests, one of a series that affirm the royal cult of the 13-year-old Pharaoh V on the first anniversary of his coronation.

Inscription Type: inscription
Inscriptions Script: demotic
Inscriptions Language:

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

June 1999 – Temporary Exhibition

Treatment date
24 June 1999

Reason for treatment
Temporary Exhibition

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

Condition
Object coated with protective treatment - identified as camisula wax, hand grease and dirt. Inscription beams traces of hard pink material, principally near edges and at corners. It closely matches the pink granite vein for colour and is covered in places by black penins ink - see annotated report. Inscription was painted with water-based black Plaka gouache, by Carol Andrews in 1982 to replace white powder. (White powder was probably modelling clay applied as a slip over entire surface: the excess would have been removed with a cloth). While powder (clay) continued to adhere areas of damage. Returns at either side have white deposit with pronounced horizontal directionality, 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 back and returns and a looser whitish material under cross-member of lid mount, but elsewhere consolidated and turned brown by camouilla wax. Text painted in white on lid return: *CAPTURED IN EGYPT BY THE BRITISH ARMY IN 1801*; on either return: "PRESENTED BY KING GEORGE III".

Treatment details
Samples were taken of pink material and printers ink in inscription; surface coating from various locations: white deposit on returns, reddish compacted clay on back and returns; loose white...
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/
Address people/hg

Projects

- ReSIST Resilience for Survivability in IST
- HELIOS

Research Areas

- Static Analysis
Concluding Remarks

- **RKBExplorer uses CRS services successfully**
  - Often 100s of URIs per person

- **British Museum uses a CRS service for Data Fusion on its web site**

- **Organisations can go ahead and build systems, then do the identity management**
  - Both internally and externally
  - Doing identity management first requires cross-organisation agreement, a delay of months to years
  - It then freezes the business processes to conform to the external requirement

- **sameAs.org turns out to be pretty useful**
  - > 10K hits/day

- **Build systems that accept the way the world is, not what you would like it to be**
Concluding Remarks

• **Consume things yourself!**
  – The only way to find out if things are any good
  – Then you might be able to:

• **Provide appropriate services in appropriate formats**
  – Consumable by “normal” systems
  – This means
    • Useful – doing useful things
    • RESTful – easy to access
    • JSON – easy to process
    • …
Endnotes

**URIs**

- [http://sameas.org/](http://sameas.org/)
- [http://sameas.org/kelle](http://sameas.org/kelle)
- [http://www.rkbexplorer.com/services/](http://www.rkbexplorer.com/services/)
- [http://apps.seme4.com/see-uk/crime/by-population/ward/OX1%203QG](http://apps.seme4.com/see-uk/crime/by-population/ward/OX1%203QG)
- [http://www.ecs.soton.ac.uk/people/dder](http://www.ecs.soton.ac.uk/people/dder)
- [http://www.dotac.info/explorer/](http://www.dotac.info/explorer/)

**You are Never Alone**

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

[http://eprints.ecs.soton.ac.uk/id/eprint/22613](http://eprints.ecs.soton.ac.uk/id/eprint/22613)