Searching Semantic Web Objects
Based on Class Hierarchies

Gong Cheng, Weiyi Ge, Honghan Wu, Yuzhong Qu
Institute of Web Science, School of Computer Science and Engineering
Southeast University, Nanjing, China

http://iws.seu.edu.cn/services/falcons/
Information Islands vs. Web of Data
Information Islands vs. Web of Data
Information Islands vs. Web of Data
Motivation

- How to find URIs?
- Linked Open Data brings …
Motivation

- How to find URIs?
- Linked Open Data brings …
- Single-point retrieval system
## Demo

http://iws.seu.edu.cn/services/falcons/

<table>
<thead>
<tr>
<th>Date</th>
<th>Document</th>
<th>Entity</th>
<th>#vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#document</td>
<td>#triple</td>
<td>#RDF_sentence</td>
</tr>
<tr>
<td>2008-03-26 17:02:24</td>
<td>9,634,401</td>
<td>356,244,433</td>
<td>285,616,738</td>
</tr>
</tbody>
</table>
Demo

Object Search  Concept Search

Beijing 2008

Supports Boolean operators, quotes, and wildcard characters.

Objects 1 - 10 of 331 for your search Beijing 2008 (0.046 seconds)

Beijing
Types: Capital, City
Labels: 北京 || Pekin || Пекин || 北京市 || Pequim || Pechino || Beijing || Pékin' || Peking || Pekín'
http://dbpedia.org/resource/Beijing - Described in 198 documents

Beijing Guoan
Types: Club
Labels: Beijing Hyundai || 北京国安” || 北京国安足球俱乐部” || Beijing Guoan
http://dbpedia.org/resource/Beijing_Guoan - Described in 36 documents

Beijing SWAT
Labels: Beijing SWAT
http://dbpedia.org/resource/Beijing_SWAT - Described in 4 documents
Demo

Object Search  Concept Search
Beijing 2008

Supports Boolean operators, quotes, and wildcard characters.

All >> Event

Act  Conference  Contest

Objects 1 - 10 of 58 for your search Beijing 2008 (0.053 seconds)

2008 Summer Olympics bids
Types: Command
Labels: 2008 Summer Olympics bids || Candidatures pour les Jeux Olympiques de 2008
http://dbpedia.org/resource/2008_Summer_Olympics_bids - Described in 5 documents

WWW2008
Types: Subject, Conference,
Labels: WWW2008
http://ontoworld.org/wiki/Special:URIResolver/WWW2008 - Described in 5 documents

2008 Summer Paralympics
Types: Game
Labels: XIII Giochi Paralimpici estivi || Paralympische Zomerspelen 2008 || 2008年夏季残疾人奥林匹克运动会" || 北京パラリンピック"
Sommer-Paralympics 2008 || 2008 Summer Paralympics
http://dbpedia.org/resource/2008_Summer_Paralympics - Described in 4 documents
Demo

Object Search  Concept Search
Beijing 2008  Search Objects

All >> Event >> Conference

Objects 1 - 1 of 1 for your search Beijing 2008 (0.072 seconds)

WWW2008
Types: Subject, Conference,
Labels: WWW2008
http://ontoworld.org/wiki/Special:URIResolver/WWW2008 - Described in 5 documents
Navigating Class Hierarchies for Query Restriction

Filtering results

2008 Summer Olympics bids
Types: Command
Labels: 2008 Summer Olympics bids || Candidatures pour les Jeux Olympiques de 2008
http://dbpedia.org/resource/2008_Summer_Olympics_bids - Described in 5 documents

WWW2008
Types: Subject, Conference,
Labels: WWW2008
http://ontoworld.org/wiki/Special:URIResolver/WWW2008 - Described in 5 documents

2008 Summer Paralympics
Types: Game
Labels: XIII Giochi Paralimpici estivi || Paralympische Zomerspelen 2008 || 2008年夏季残疾人奥林匹克运动会” || 北京パラリンピック”
Sommer-Paralympics 2008 || 2008 Summer Paralympics
http://dbpedia.org/resource/2008_Summer_Paralympics - Described in 4 documents

Recommending subclasses
Combined Inverted Index

- Query: `<terms, classes>`
Provenance of Typing Information

Heuristics

- Dereference document
- Namespace document
- Documents on the same host
- Other documents
Class-inclusion Reasoning

Diagram showing a hierarchy of classes:

- **swrc:Person**
  - **swrc:Employee**
    - **swrc:AcademicStaff**
  - **swrc:Student**
    - **swrc:Manager**
    - **swrc:Graduate**

- **Terms**
  - professor
  - tennis

- **Classes**
  - lws:Qu
  - cse:Li
  - lws:Zhai
  - lws:Gao
  - lws:Wu
  - lws:Zhang
  - swrc:AcademicStaff
  - swrc:Manager
  - swrc:Graduate

- **swrc:Graduate**
  - swrc:Manager

Gong Cheng  gcheng@seu.edu.cn
Class-inclusion Reasoning (cont.)
Which axioms would be accepted by the reasoning engine?

- The authorized description of a vocabulary is allowed to reuse the classes from other vocabularies but **CANNOT** further constrain their meaning.
Class-inclusion Reasoning (cont.)

Which axioms would be accepted by the reasoning engine?

- The authorized description of a vocabulary is allowed to reuse the classes from other vocabularies but **CANNOT** further constrain their meaning.
Recommending Subclasses

- If a subclass covers more results, it will be more likely to be recommended.
Recommending Subclasses

- If a subclass covers more results, it will be more likely to be recommended.

- Algorithm
  - Iterate over the first 1,000 results to collect classes.
  - Rank classes.
  - Select top $K$ classes, s.t.:
    - Each selected class must be a strict subclass of a current class.
    - The class inclusion relation does NOT hold between any pair of the selected classes.
Example

\[ K = 2 \]

<table>
<thead>
<tr>
<th>Name</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>iws:Qu</td>
<td>swrc:AcademicStaff</td>
</tr>
<tr>
<td></td>
<td>swrc:Employee</td>
</tr>
<tr>
<td></td>
<td>swrc:Person</td>
</tr>
<tr>
<td>cse:Li</td>
<td>swrc:Manager</td>
</tr>
<tr>
<td></td>
<td>swrc:Employee</td>
</tr>
<tr>
<td></td>
<td>swrc:Person</td>
</tr>
<tr>
<td>iws:Zhai</td>
<td>swrc:AcademicStaff</td>
</tr>
<tr>
<td></td>
<td>swrc:Employee</td>
</tr>
<tr>
<td></td>
<td>swrc:Person</td>
</tr>
<tr>
<td>iws:Gao</td>
<td>swrc:AcademicStaff</td>
</tr>
<tr>
<td></td>
<td>swrc:Employee</td>
</tr>
<tr>
<td></td>
<td>swrc:Person</td>
</tr>
<tr>
<td>iws:Wu</td>
<td>swrc:Graduate</td>
</tr>
<tr>
<td></td>
<td>swrc:Student</td>
</tr>
<tr>
<td></td>
<td>swrc:Person</td>
</tr>
<tr>
<td>iws:Zhang</td>
<td>swrc:Graduate</td>
</tr>
<tr>
<td></td>
<td>swrc:Student</td>
</tr>
<tr>
<td></td>
<td>swrc:Person</td>
</tr>
</tbody>
</table>

Current class:
- swrc:Person 6
- swrc:Employee 4
- swrc:Student 2
- swrc:Manager 1
- swrc:Graduate 2
- swrc:AcademicStaff 3
### Summarizing SW Objects

<table>
<thead>
<tr>
<th>General</th>
<th>DOC</th>
<th>WOT</th>
<th>CONTACT</th>
<th>DG</th>
<th>DOAP</th>
<th>FOAF</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Tim Berners-Lee] described in FOAF: <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/</a></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**phone** --> 

**sha1sum of a personal mailbox URI name**

965c47c5a70db7407210cef6e4e6f5374a525c5c

**personal mailbox** --> 

mailto:timbl@w3.org

**based near** --> 

+ [G2]

+ [G1]

**Given name**

Timothy

**nickname**

TimBL

timbl

**family name**

Berners-Lee

**member** --> 

DIG (http://dig.csail.mit.edu/data#DIG)

**name**

Tim Berners Lee
Tim Berners-Lee
Timothy Berners-Lee

**openid** -->

Tim Berners-Lee (http://www.w3.org/People/Berners-Lee)
Mapping Terms to SW Objects

Expanding textual description of SW objects

- wiki:Category-3AConference
  - rdfs:label: Conference
  - rdf:type: wiki:WWW2008
- wiki:Beijing
  - property:Has_location_city
- wiki:WWW2008
  - property:Title: 17th International World Wide Web Conference
  - rdfs:label: WWW2008
Thank You

Welcome to Nanjing!