Learning Linked Data

Introducing a Competency Framework & Related Learning resources

Marcia Lei Zeng
Kent State University

Northern Ohio Technical Services Librarians (NOTSL) Fall Meeting
December 11, 2017, Parma, Ohio.
Outline

A. The Impact of Linked Data to Library professionals
   1) Refresh a few concepts
      • based on our 2013 NOTSL discussion of “Linked Data and LAMs Why and What?”
   2) KOS (knowledge organization systems) trending
      • Demo: How to benefit from the LOD KOS products

B. Introducing a Competency Framework & Related Learning resources
1) Refresh a few concepts

Linked Data Principles (Berners-Lee 2006)

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)
4. Include links to other URIs so that they can discover more things

http://www.w3.org/DesignIssues/LinkedData.html

Update: IRI

URI (Uniform Resource Identifier)
IRI (Internationalized Resource Identifier)

http://dbpedia.org/page/Himalayas

Using URI/IRI to represent an entity of resource:
-- one of the basic solutions to providing machine-processable, disambiguated data.
1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)
4. Include links to other URIs so that they can discover more things

<subject> <predicate> <object>

Resource: Pride and Prejudice

Property: creator

Property value: Jane Austen

"Statement" with dc:
- OCLC:246662790 dc:title: Pride and prejudice
- OCLC:246662790 dc:creator: Austen, Jane
- OCLC:246662790 dc:subject: Social classes -Fiction

"Statement" with dcterms:
1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)
4. Include links to other URIs so that they can discover more things

Case: WorldCat

The changing concepts (seeing from the content)

- From "Web of Documents" to "Web of Data"
- From linking strings to linking things

(cont.) Linked Data Principles

<table>
<thead>
<tr>
<th>URI</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://worldcat.org/entity/work/id/25347688">http://worldcat.org/entity/work/id/25347688</a></td>
<td>Cataloging and classification: an introduction</td>
</tr>
<tr>
<td><a href="http://worldcat.org/entity/person/id/2631227899">http://worldcat.org/entity/person/id/2631227899</a></td>
<td>Lois Mai Chan</td>
</tr>
</tbody>
</table>

Linking out: Linking THINGS, not strings.
Access through the linked things.

http://www.worldcat.org/oclc/922220005
1) Refresh a few concepts

The changing concepts (seeing from the content)

- From "Web of Documents" to "Web of Data"
- From linking strings to linking things

Go to [http://dbpedia.org/page/Lois_Mai_Chan](http://dbpedia.org/page/Lois_Mai_Chan) and follow knownFor and about

Linking THINGS, not strings. Access through the linked things.
1) Refresh a few concepts

**The changing concepts**
(seeing from the content)

- From "Web of Documents" to "Web of Data"
- From linking **strings** to linking **things**

(seeing from the results)

- From "On the Web" to "Of the Web"

**Linking THINGS, not strings.**
Access through the linked things.
Enhancing semantic consistency of data through shared, unconventional mashup activities

Update: mashup culture

https://en.wikipedia.org/wiki/Frank_Lloyd_Wright
Update: mashup culture

Wikidata – as an Authority base/ Knowledge Base

- Defined and controlled
- Mapped to other Vocab IDs
- Linked to examples
Case: FAST

Source: extracted screenshots (2017-07-12)
From http://fast.oclc.org/searchfast/
John F. Kennedy’s entry in FAST is enriched with other sources.

- The **DBpedia** identifiers allow FAST terms to include detailed information that is usually excluded in authority records.

- The **VIAF URI** allows FAST terms to take advantage of all of the various string values included in **VIAF** without having to manually include the values in the RDF triples for the specific term.

Source: extracted screenshots (2017-07-12) at http://experimental.worldcat.org/fast/35588/rdf.xml
The GeoNames data is used to power MapFAST, which is a Google Maps mash-up.
Swissbib, a provider for bibliographic data in Switzerland
http://linked.swissbib.ch

- 30,773 links to DBpedia
- 20,714 links to VIAF
- high precision values
- generated in reasonable expenditures of time

Outline

A. The Impact of Linked Data to Library professionals
   1) Refresh a few concepts
      • based on our 2013 NOTSL discussion of “Linked Data and LAMs Why and What?”
   2) KOS (knowledge organization systems) trending
      • Demo: How to benefit from the LOD KOS products

B. Introducing a Competency Framework & Related Learning resources
In the BARTOC registry
(thesaurus, ontology, classification)
KOS vocabularies registered:
-(2016-05): 1836
-(2017-12-09): 2,743
RDF: 315

2) KOS trending

In the Datahub

LOD KOS registered
-(2016-03): 1251 found
-(2017-08): 1662 found
with tags “thesaurus”, “classification”, “taxonomy”, “ontology”, “terminology”
(Note: some are tagged with multiple categories. Some have multiple editions.)

http://bartoc.org/
https://datahub.io/
Growth of the KOS for the information and knowledge organization tasks

Classification, Subject headings
- LIS professionals
- I&A database producers

+ Thesaurus
- LIS professionals
- I&A database producers

++ Ontology, knowledge base
- LIS professionals
- I&A database producers
- Anyone
Growth of the KOS for the information and knowledge organization tasks

Conventional KOS have always been quick adapters of new technologies in their publishing venues and applications.

The release of a LOD KOS product represents a turning point for the producer or provider of a vocabulary.
Demo: How to benefit from the LOD KOS products

• Advanced uses of thesauri that are released with multiple services: Websites, APIs, and SPARQL endpoints

  See another set of ppt
Advanced uses of thesauri that are released with multiple services: Websites, APIs, and SPARQL endpoints

1. Learn through TGN Website http://www.getty.edu/research/tools/vocabularies/tgn/ (Anyone can do!)

2. Following these geographic places located on the Silk Road, using the geo-coordinators TGN provided, get them on the map (through TGN’s API). (Need someone who can play with API and write a little bit Java.)

   • The APIs are available to any institution having a login, which may be obtained by writing to vocab@getty.edu. See details in the Web Services User's Instructions (PDF).

3. Find certain place types around the Silk Road through a LOD Sparql Query platform http://vocab.getty.edu/ (Anyone can use the template to query; follow our simple demos; knowledge of SPARQL queries will be ideal.)
(Cont.) Demo: How to benefit from the LOD KOS products

Creating new KOS vocabularies for a particular project’s products by extracting the components from a comprehensive KOS vocabulary.

Steps:
1. Go to Getty Vocab LOD SPARQL Endpoint: http://vocab.getty.edu/sparql
2. Choose ‘Queries’.
3. Choose "Descendants of a Given Parent" from the template, click. → Now, the template's text will show on the right.
4. Click ‘SPARQL” to get the query text up.
5. Use this ID in the query, send the query.

6. Get the dataset, download.
(Cont.) Demo: How to benefit from the LOD KOS products

Name authorities offer foundational structured data for network analyses.

At the same query templates page
http://vocab.getty.edu/queries

Find the section for ULAN.
- There are many interesting query examples.
Steps: (1) go to 5.2; (2) click on that SPARQL sign for 5.2; (3) replace the ID of the person you want to find. Note there are two IDs to replace. (4) Submit.

Query: Find associative relationships of ulan:500020307 Wright, Frank Lloyd (American architect, 1867-1959); showing relationship type, associated persons, each person’s preferred name, preferred display biography, and other notes.
<table>
<thead>
<tr>
<th>rel</th>
<th>x</th>
<th>name</th>
<th>bio</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>gvp:ulan1000_related_to</td>
<td>ulan:500077136</td>
<td>Sullivan, Francis</td>
<td>Canadian architect and draftsman, 1882-1929</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1106_apprentice_was</td>
<td>ulan:500035255</td>
<td>Neutra, Richard</td>
<td>American architect, 1892-1970, born in Austria</td>
<td>Taliesin, Wisconsin, fall of 1924</td>
</tr>
<tr>
<td>gvp:ulan1106_apprentice_was</td>
<td>ulan:500255776</td>
<td>Robinson, Harry Franklin</td>
<td>American architect and draftsman, 1883-1959</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1106_apprentice_was</td>
<td>ulan:500249945</td>
<td>Smith, Tony</td>
<td>American sculptor, architect, and painter, 1912-1980</td>
<td>1938-1940</td>
</tr>
<tr>
<td>gvp:ulan1218_employee_of</td>
<td>ulan:50001446</td>
<td>Soleri, Paolo</td>
<td>American architect born in Italy, 1919-2013</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1202_patron_was</td>
<td>ulan:500071769</td>
<td>Giannini and Hilgart</td>
<td>American glass stainiers, established 1899</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1302_associate_of</td>
<td>ulan:500114125</td>
<td>Niedecken-Walbridge Co.</td>
<td>American interior design firm, established 1907</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1303_collaborated_with</td>
<td>ulan:500045662</td>
<td>Niedecken, George M.</td>
<td>American interior decorator, 1878-1945</td>
<td>1904-1918</td>
</tr>
<tr>
<td>gvp:ulan1303_collaborated_with</td>
<td>ulan:50040695</td>
<td>Green, Aaron G.</td>
<td>American architect, 1917-2001</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1308_assisted_by</td>
<td>ulan:50028814</td>
<td>Endo, Arata</td>
<td>Japanese architect, 1889-1961</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1501_sibling_of</td>
<td>ulan:500116874</td>
<td>Barney, Maginel Wright</td>
<td>American illustrator, 1881-1966</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1511_child_of</td>
<td>ulan:50003241</td>
<td>Wright, Anna Lloyd</td>
<td>American teacher, died 1923</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1512_child_of</td>
<td>ulan:500035777</td>
<td>Wright, John Lloyd</td>
<td>American architect and designer, 1892-1972</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1512_parent_of</td>
<td>ulan:500038812</td>
<td>Wright, Lloyd</td>
<td>American architect and scenographer, 1890-1978</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1514_ancestor_of</td>
<td>ulan:500030393</td>
<td>Wright, Eric Lloyd</td>
<td>American architect, born 1929</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1541_spouse_of</td>
<td>ulan:500218907</td>
<td>Wright, Catherine Tobin</td>
<td>American socialite, 1871-1959</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1541_spouse_of</td>
<td>ulan:500260870</td>
<td>Wright, Olivianna Lloyd</td>
<td>American socialite, 1900-1985</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1541_spouse_of</td>
<td>ulan:500281330</td>
<td>Oak Park Studio</td>
<td>American architectural firm, established ca. 1996, dissolved 1909</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan1541_spouse_of</td>
<td>ulan:50023286</td>
<td>Oak Park Studio</td>
<td>American architectural firm, established ca. 1896, dissolved 1909</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan2781_dedicated_to</td>
<td>ulan:5000309156</td>
<td>Taliesin</td>
<td>American repository, Spring Green, contemporary</td>
<td>-</td>
</tr>
<tr>
<td>gvp:ulan2781_dedicated_to</td>
<td>ulan:500308851</td>
<td>Taliesin West</td>
<td>American repository, Scottsdale, contemporary</td>
<td>-</td>
</tr>
</tbody>
</table>
Outline

A. The Impact of Linked Data to Library professionals
   1) Refresh a few concepts
      • based on our 2013 NOTSL discussion of “Linked Data and LAMs Why and What?”
   2) KOS (knowledge organization systems) trending
      • Demo: How to benefit from the LOD KOS products

B. Introducing a Competency Framework & Related Learning resources