Minnesota Zine Linked Data
Working Group:
The First Project

Upper Midwest Users Group (Ex Libris)
Oct. 28, 2013
Roseville, MN
What we talk about when we talk about linked data
A formal statement of concepts used within a knowledge domain, and the relationships between those concepts

- Genealogical ontologies
- Taxonomic ontologies (“King Philip Came Over From Greece Saturday”)
- FRBR
- Dublin Core
Principles of linked data:

1. Use URIs for the names of things
2. Use http URIs so they can be looked up
3. When someone looks up a URI, provide useful information
4. Include links to other URIs
<table>
<thead>
<tr>
<th>Detail</th>
<th>Properties</th>
<th>History</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferred Label</td>
<td>Zombies</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>English</td>
<td></td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://metadataregistry.org/souch/1841">http://metadataregistry.org/souch/1841</a></td>
<td></td>
</tr>
<tr>
<td>Status</td>
<td>Published</td>
<td></td>
</tr>
</tbody>
</table>

### Properties

<table>
<thead>
<tr>
<th>metasheet</th>
<th>Label</th>
<th>Preferred Label</th>
<th>Language</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zombies</td>
<td>Zombies</td>
<td>English</td>
<td>Published</td>
<td></td>
</tr>
</tbody>
</table>
Why?
Twin Cities Zine Linked Data Working Group

An explanation of the project
MCTC Library Zine Collection

Zines: a self publication, motivated by a desire for self expression, not for profit.
Anchor Archive Subject Thesaurus

Developed for a special collection of zines in Halifax, Canada

Zine Thesaurus of Subject Terms

ABANDONED BUILDINGS
- RT
  - GHOST TOWNS
  - URBAN DECAY

ABLEISM
- RT
  - DISABILITIES
  - PHYSICAL DISABILITIES

ABORIGINAL PEOPLES
- USE
  - INDIGENOUS PEOPLES
Creating a Linkable Subject Authority

Open Metadata Registry allows people and organizations to “identify, declare or publish” metadata schemas
Creating a Linkable Subject Authority

Our working group rebuilt the Anchor Archive Subject Thesaurus in the Open Metadata Registry.
Creating a Linkable Subject Authority

Open Metadata Registry allows us to map relationships within the Subject Thesaurus

- Broader
- Narrower
- Preferred ("Alternative")
- Scope Notes
Creating a Linkable Subject Authority

Open Metadata Registry also helps make these metadata schemas linkable by creating a SKOS framework for controlled vocabularies or thesauri.

This is a SKOS file that you can download and save somewhere.
Open Metadata Registry

• **Pros**
  - Easy to use
  - Web based
  - Creates RDF links

• **Cons**
  - Labor intensive
  - Takes time
  - No spell checks
  - No de-duplication
Open Refine

- Open Refine is a free open source tool for cleaning up large metadata sets. It can also be used to create linkages between metadata sets.

  • freeyourmetadata.org has instructions for using Open Refine to clean up metadata and link it to other data sets.
Using Open Refine to Link

Step 1: Place your metadata in Open Refine.

Step 2: Make sure you have the RDF extension.
Using Open Refine to Link

Step 3: Using the RDF extension, you set up a reconciliation service.

- I will need to use based on RDF file....
Using Open Refine to Link

- Name the reconciliation service
- Upload your SKOS file from Open Metadata Registry
- make sure you check SKOS in the label properties area
Using Open Refine to Link

Step 4: Select the column you want to reconcile, and navigate to start reconciling...

- I want to match my Subject Headings
Using Open Refine to Link

Step 5: Select your newly created reconciliation service

- In my case, it’s AAST (Anchor Archive Subject Thesaurus)
Using Open Refine to Link

Step 6: Open Refine will reconcile your metadata with the linked data of you choose.

- If you have lots of metadata, this is the “Go get some Coffee” step
Using Open Refine to Link

Step 7: Select the best matches

- The blue headings are linked to the Anchor Archive Subject Thesaurus in the Open Metadata Registry
Using Open Refine to Link

• You can match a single subject heading at a time
Using Open Refine to Link

- Or you can match multiple similar cells at once
Thoughts on the Linked Data Project

• MCTC Zine Collection metadata is not ready for prime time

• Open Metadata Registry can also produce SPARQL endpoints
What Can We Do Now?

Some of the uses to which it could be put
Basically what we’ve done is create URIs for these subject headings, giving them a permanent central location on the web.
But that doesn’t really mean anything unless someone actually uses the subject headings within a collection.
One thing we could do is enter a collection of zines, such as the one at MCTC, into an online database and use the URIs for the subjects that we entered into the Open Metadata Registry.
One database possibility to explore is Freebase.com, "a community-curated database of well-known people, places and things."
“Freebase is a database of things”
"All things in Freebase have strong identifiers"
The data in Freebase is harvested from certain websites (like Wikipedia) or is directly entered by individuals like you or me.
Things in Freebase can also be entered in RDF form, which means that the content in Freebase can be used as linked data.
If we chose to, we could take MCTC’s zine collection and enter it in Freebase.

In the metadata we create for each zine in Freebase, we could link to the subject URIs that we created in the Open Metadata Registry.
Each zine entered in Freebase would become linked data that could be used on other websites. Each zine would also link back (through the subjects) to more zines on the same topic.
Or at least, that's our theory... 😊
We could use the vocabulary data to provide other kinds of support for searching

- Non-preferred terms
- Broader, narrower, and related terms
- Scope notes
- Synonym rings
Another possibility would be to map terms in our Zine vocabulary to other vocabularies.
Public transit
(http://metadataregistry.org/uri/aast1012)

Skos:closeMatch

Local transit
(http://id.loc.gov/authorities/subjects/sh85077945)
Another way to let searchers connect with our resources
Similar Projects
The project to publish all four of the Getty vocabularies as Linked Open Data has begun in earnest. It is anticipated that the data will be published under the ODC_BY 1.0 license.
Publishing Authorities and Vocabularies

LC Linked Data Service
 Authorities and Vocabularies

Search

Enter Keyword or Phrase

Available Datasets

The Linked Data Service provides access to commonly found standards and vocabularies promulgated by the Library of Congress. This includes data values and the controlled vocabularies that house them. The following are currently offered as part of this service:

» LC Subject Headings
» LC Name Authority File
» LC Classification
» LC Children's Subject Headings
» LCSH Conr/Form Terms
» Thesaurus for Graphic Materials
» Cultural Heritage Organizations

» MARC Relaters
» MARC Countries
» MARC Geographic Areas
» MARC Languages
» ISO639-1 Languages
» ISO639-2 Languages
» ISO639-5 Languages
» Extended Date/Time Format

**Please Note:** LC Classification entries are not included in general search results. You must explicitly select LC Classification in order to search the scheme. This is temporary while the impact of adding LCC to the current system is better understood.
Project: Location LODer

Dublin is the capital and most populous city of Ireland. The English name for the city is derived from the Irish name Dubhlinn, meaning "black pool". Dublin is situated near the midpoint of Ireland's east coast, at the mouth of the River Liffey and the centre of the Dublin Region. Originally founded as a Viking settlement, it evolved into the Kingdom of Dublin and became the island's principal city following the Norman invasion. The city expanded rapidly from the 17th century. It was briefly the second largest city in the British Empire and the 14th largest in Europe. Dublin entered a period of stagnation following the Act of Union of 1800, but it remained the economic centre for most of the island.
Project: Linked Jazz
What we learned