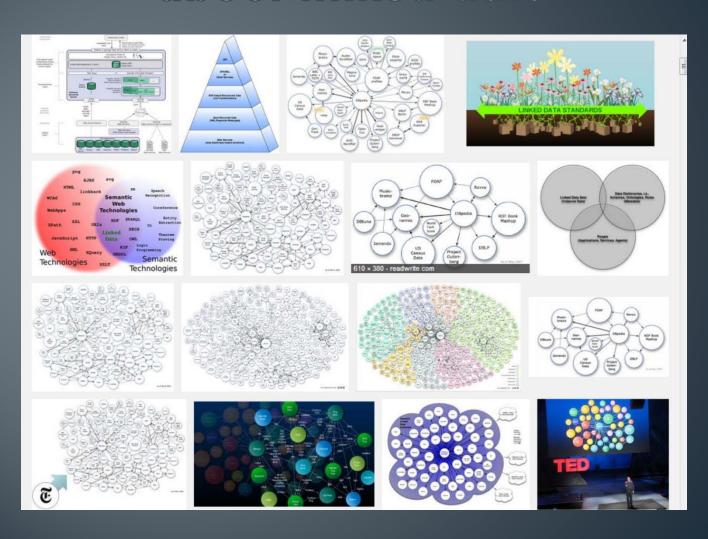
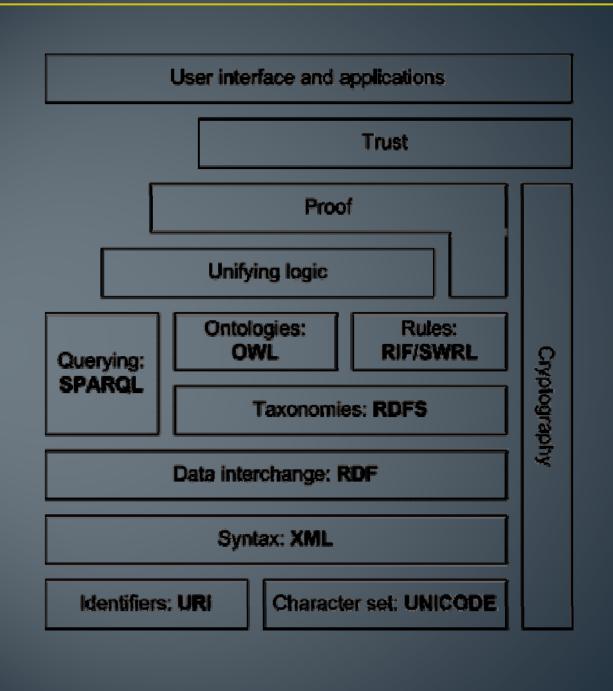


# What we talk about when we talk about linked data

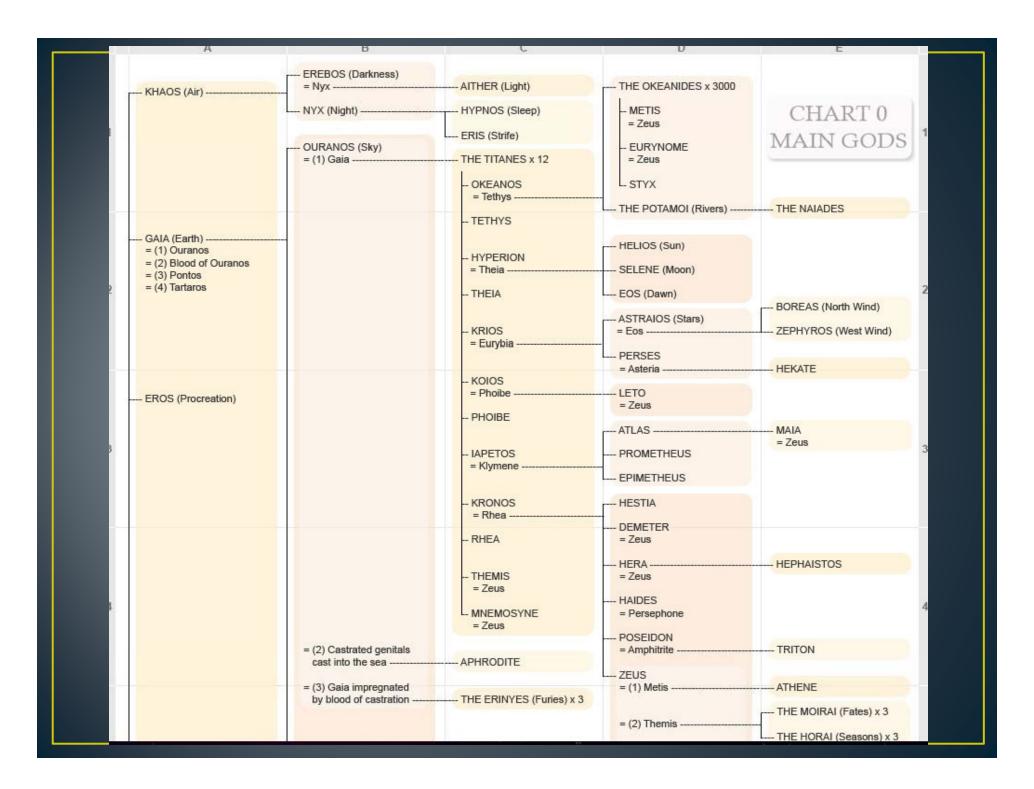




### How: Ontologies

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



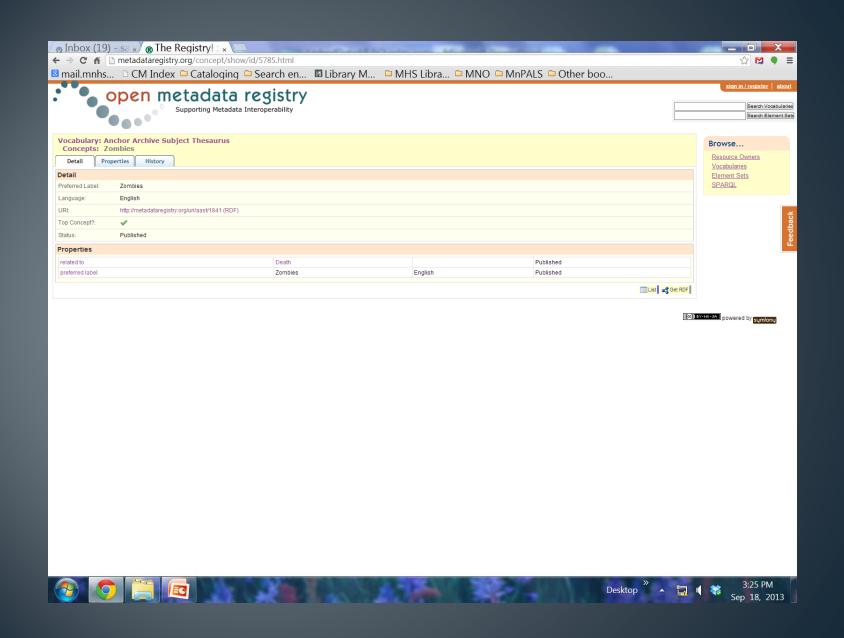
# How: The names of things, structured data, and URIs

#### Principles of linked data:

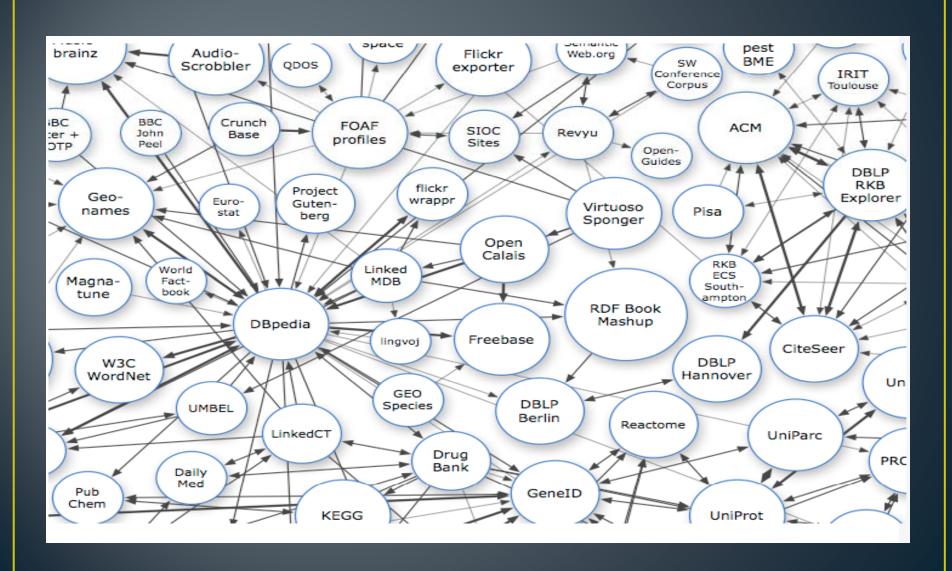
- 1. Use URIs for the names of things
- 2. Use http URIs so they can be looked up
- 3. When someone looks up a URI, provide useful information
- 4. Include links to other URIs

Berners-Lee, Tim. Linked Data. 2006-07-27.

http://owl.english.purdue.edu/owl/resource/747/08/. Accessed September 18, 2013.



# 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.







#### What is a zine?

A zine is a selfpublication, motivated by a desire for selfexpression, not for profit.

#### How big is the MCTC Zine Collection?

Approximately 5,000 zines strong with 1,000 zines to be added!

#### Is the collection open to the public?

Yes, you can stop by and read them at the library, you can't sheck them out and take them home with you.

#### Why do you collect zines?

We actively develop and collect alternative press materials, zines are just one alternative publication we collect.

Donations always welcome: MCTC Library Zine Collection 1501 Hennepin Avenue Minneapolis, MN 55403

Contact for our Zine Collection: Elissah Becknell Elissah.becknell@minneapolis.edu



### 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

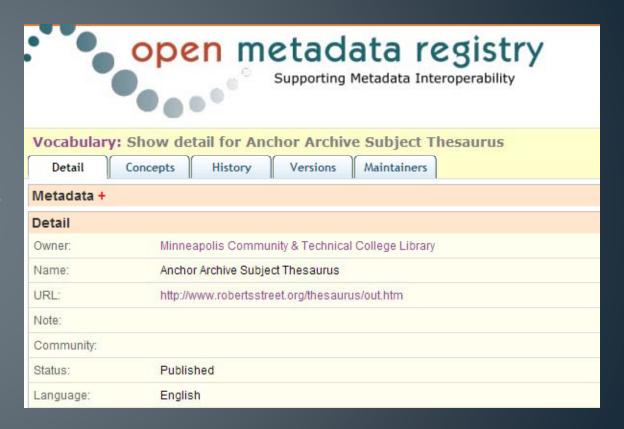
#### ABORIGINAL PEOPLES

- USE
  - INDIGENOUS PEOPLES

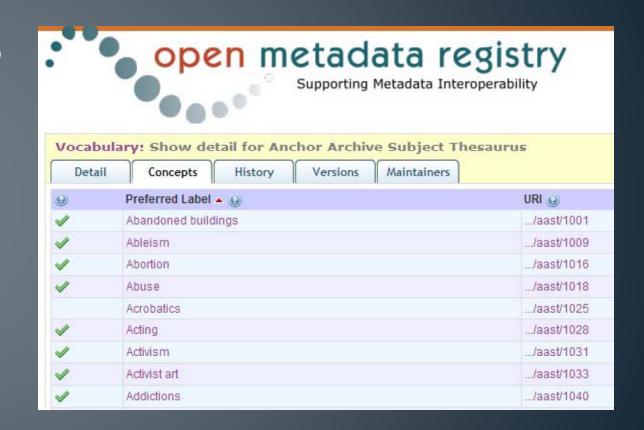
#### Anchor Archive Zine Library



Open Metadata
Registry allows
people and
organizations to
"identify, declare or
publish" metadata
schemas



Our working group rebuilt the Anchor Archive Subject Thesaurus in the Open Metadata Registry

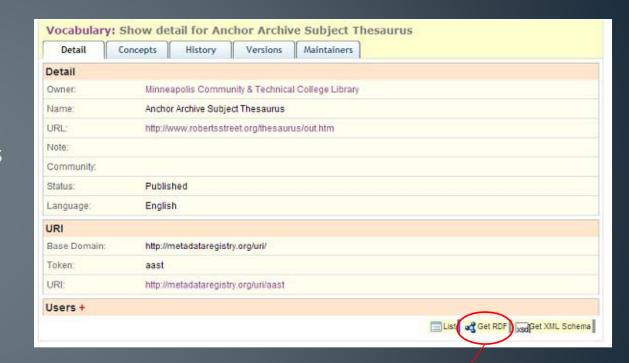


Open Metadata
Registry allows us to
map relationships
within the Subject
Thesaurus

- Broader
- Narrower
- Preferred ("Alternative")
- Scope Notes



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



Vocabulary: Show detail for Anchor Archive Subject Thesaurus	
Detail	Concepts History Versions Maintainers
0	Preferred Label @
4	Abandoned buildings
1	Urban decay
	Theatre
1	Colonization
1	Breast implants
4	Ableism
4	Public spaces
* * * * * * * * * * * * * * * * * * *	Goth
1	Breast self-examination
	Public transit
<b>4</b>	Abortion
4	Medical system
4 4 4	Abuse
4	Breastfeeding
4	Breasts
1	Graffiti
<b>4</b>	Disabilities
	Acrobatics
<b>V</b>	Grammar

#### 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.

#### **OpenRefine**

A free, open source, power tool for working with messy data

### Free Your Metadata

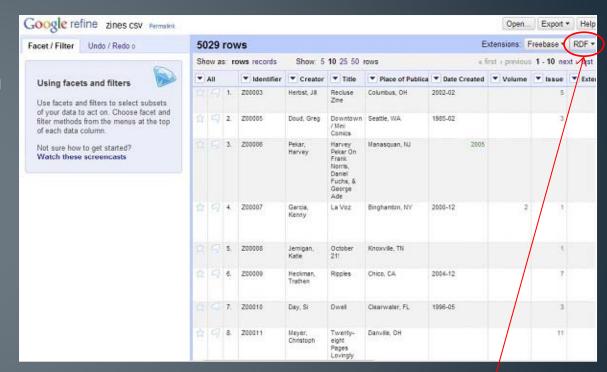
Learn how to get more value out of metadata easily



• <u>freeyourmetadata.org</u> has instructions for using Open Refine to clean up metadata and link it to other data sets.

Step 1: Place your metadata in Open Refine.

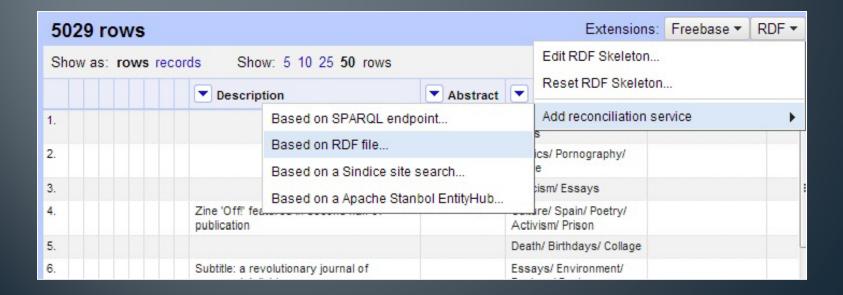
Step 2: Make sure you have the RDF extension.



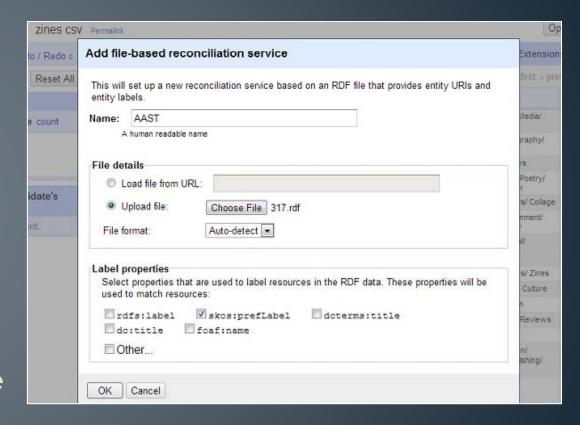
This RDF extension

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

• I will need to use based on RDF file....

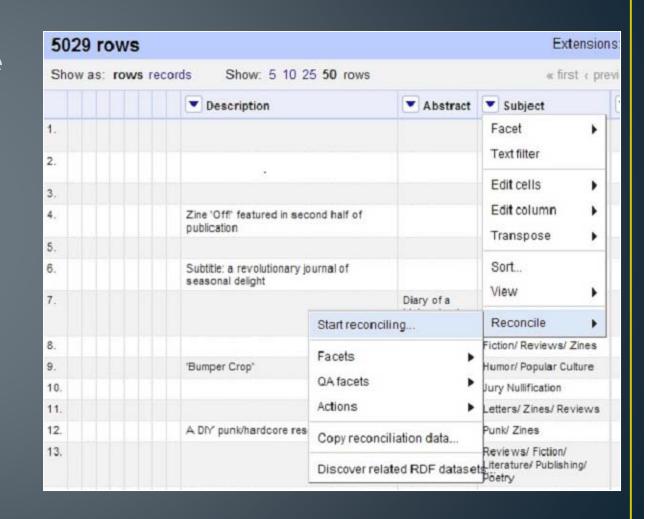


- Name the reconciliation service
- Upload your SKOS file from Open Metadata Registry
- make sure you check SKOS in the label properties area



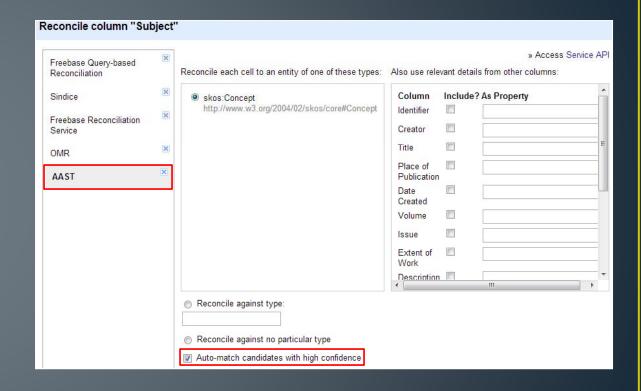
Step 4: Select the column you want to reconcile, and navigate to start reconciling...

I want to match my Subject Headings



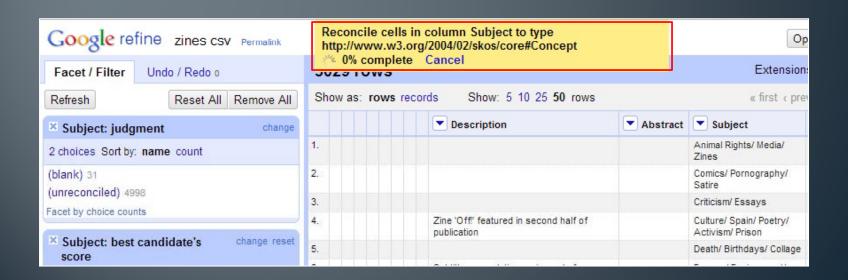
Step 5: Select your newly created reconciliation service

In my case, it's
 AAST (Anchor
 Archive Subject
 Thesaurus)



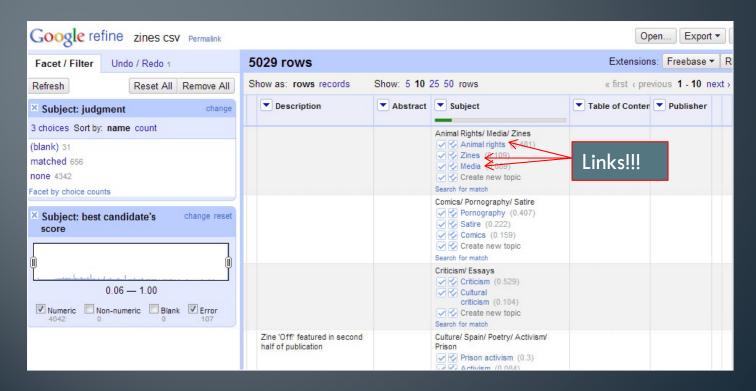
Step 6: Open Refine will reconcile your metadata with the linked data of you choose.

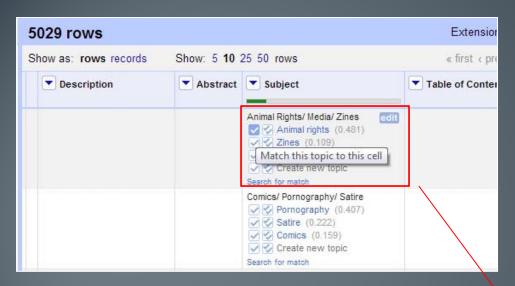
 If you have lots of metadata, this is the "Go get some Coffee" step



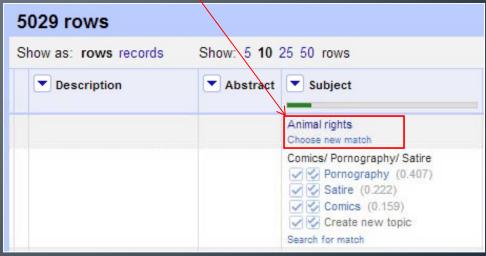
#### Step 7: Select the best matches

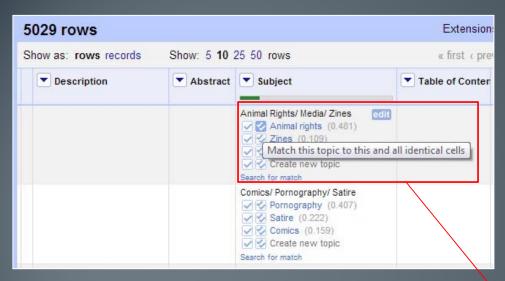
The blue headings are linked to the Anchor Archive Subject
 Thesaurus in the Open Metadata Registry



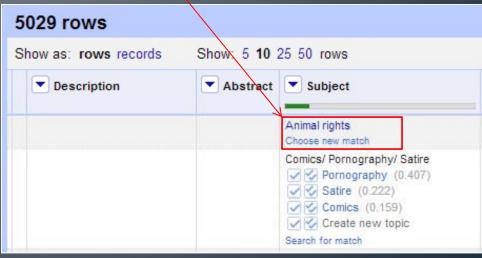


 You can match a single subject heading at a time





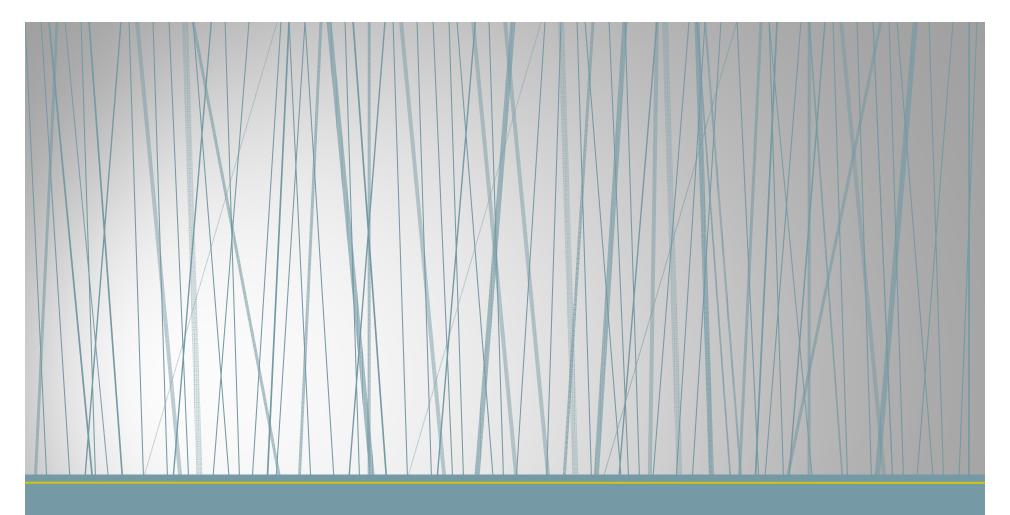
 Or you can match multiple similar cells at once



### Thoughts on the Linked Data Project

 MCTC Zine Collection metadata is not ready for prime time

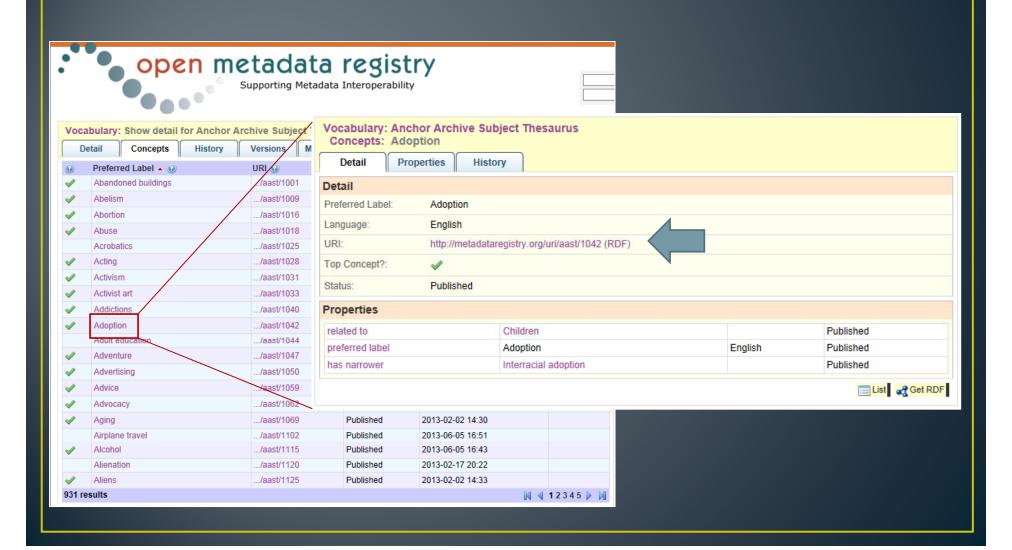
Open Metadata Registry can also produce SPARQL endpoints



What Can We Do Now?

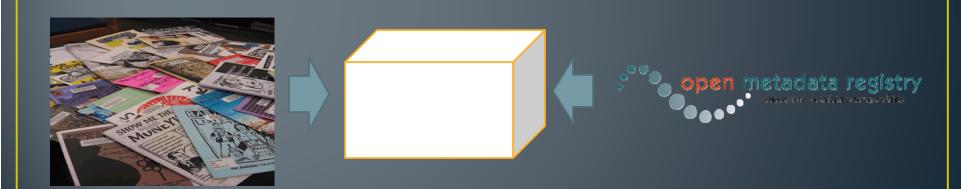
Some of the uses to which it could be put

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.

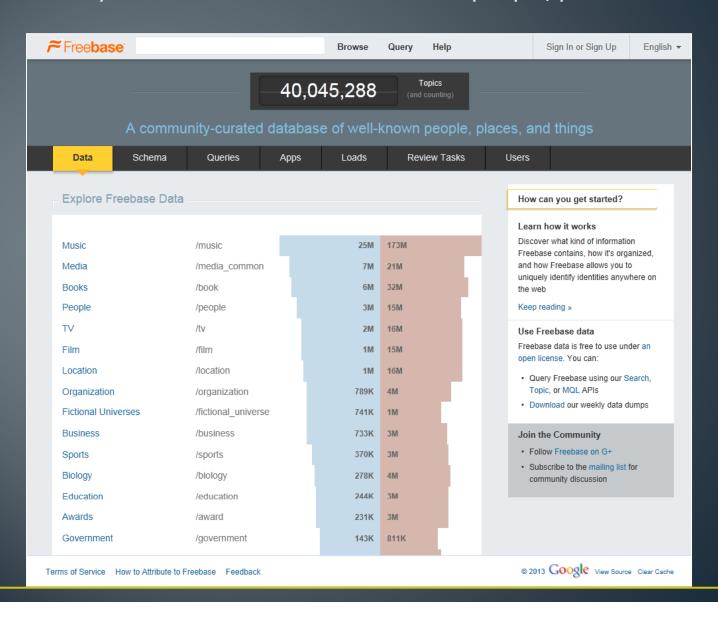
One thing we could do is enter a collection of zines, such as the one at MCTC, into an online database and use the URIs for the subjects that we entered into the Open Metadata Registry.



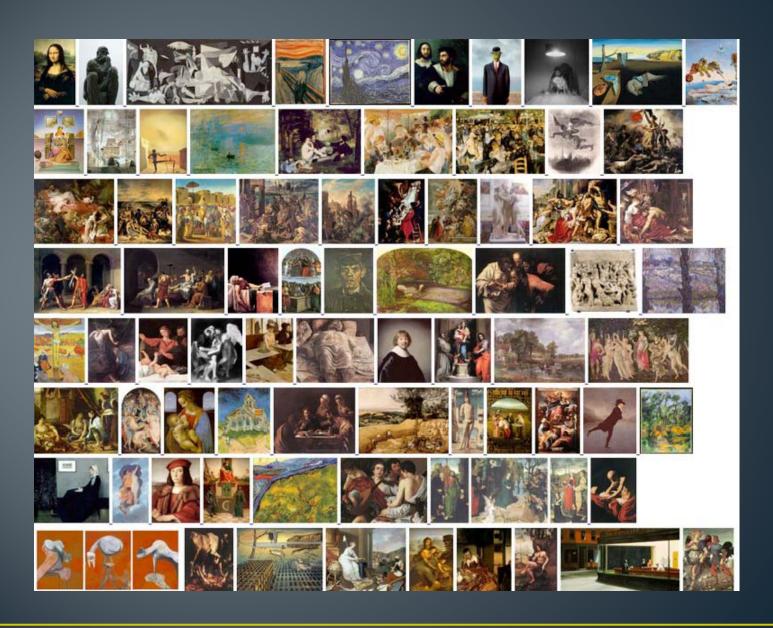
Zine collection

#### 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"















/en/mona lisa /en/the thinker





/en/the scream /en/the starry night /en/self portrait with a friend /en/the son of man







/en/untitled man and mirror /en/the persistence of memory /en/dream caused by the flight of a bee around a pomegranate a second before awakening







/en/the\_madonna of port\_lligat /en/the\_discovery\_of america by christopher\_columbus /en/the\_ghost\_of\_vermeer\_of\_delft\_which\_can\_be\_used\_as\_a\_table









/en/impression\_sunrise /en/the\_luncheon\_on\_the\_grass /en/luncheon\_of\_the\_boating\_party /en/bal\_au\_moulin\_de\_la\_galette\_montmartre

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.







Zine + metadata (includes a zine URI)



Subject URI:
Urban 

Exploration





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

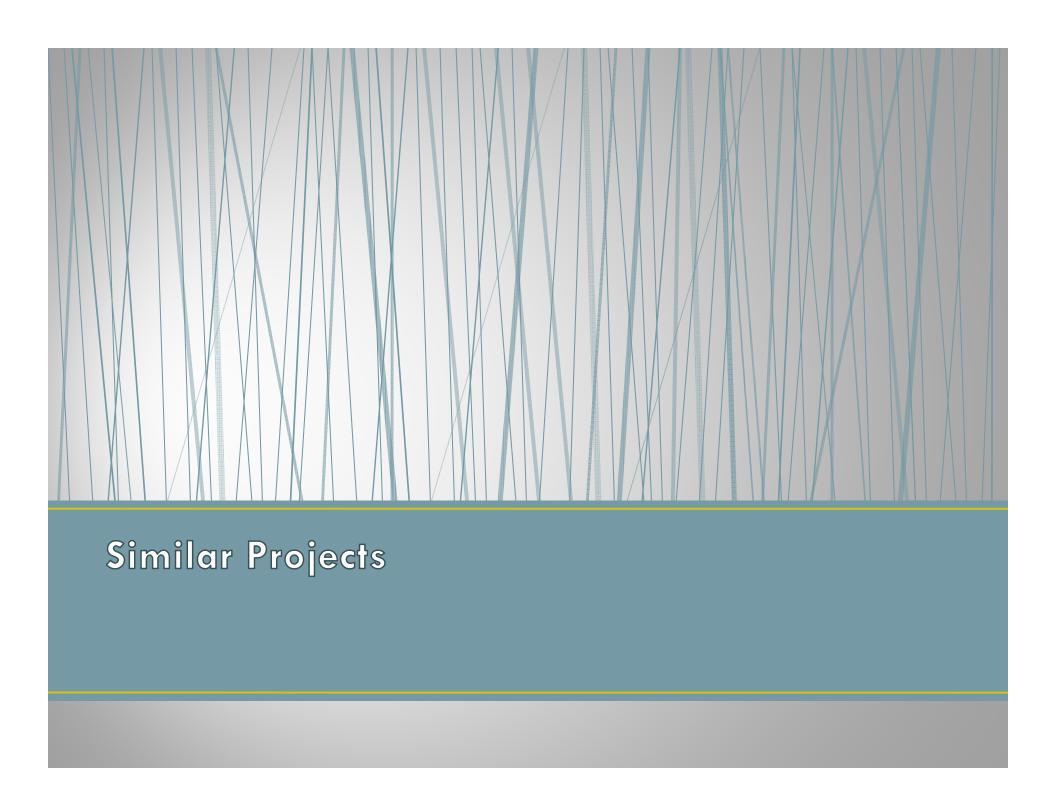
# Public transit (http://metadataregistry.org/uri/aast1012)

Skos:closeMatch

Local transit (http://id.loc.gov/authorities/subjects/sh85077945)

# Another way to let searchers connect with our resources

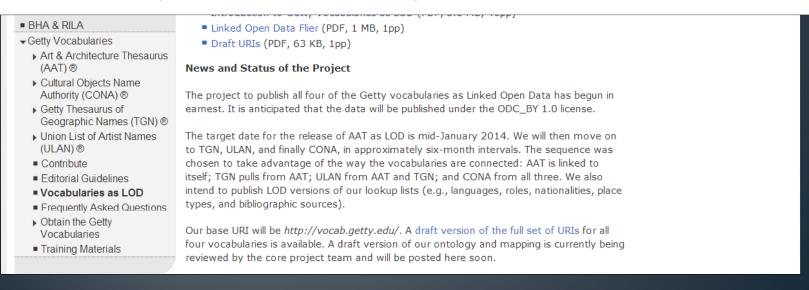




## Publishing Vocabularies



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

#### LIBRARY OF CONGRESS LINKED DATA SERVICE

- Data Service
- About
- Main Dataset Descriptions
- Preservation Dataset Descriptions
- Search
- Download
- Technical Center
- Contact Us

#### LC Linked Data Service Authorities and Vocabularies





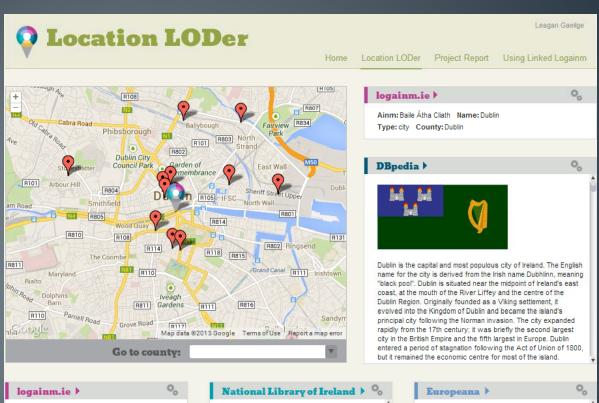
#### **Available Datasets**

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
- > LC Genre/Form Terms
- > Thesaurus for Graphic Materials
- > Cultural Heritage Organizations

- > MARC Relators
- > MARC Countries
- > MARC Geographic Areas
- > MARC Languages
- > ISO639-1 Languages
- > ISO639-2 Languages
- > ISO639-5 Languages
- > Extended Date/Time Format

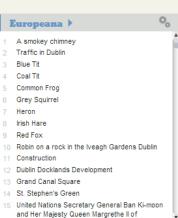
### Project: Location LODer







A survey of lands and premises lying between Aughrim St and Prussia St done for General Hawkshaw. By John Longfield 1815. Scale 40 feet to an Inch. Tenants names show on



## Project: Linked Jazz

