Create Microthesauri and other datasets from the Getty LOD Vocabularies

Marcia Lei Zeng mzeng@kent.edu

Museums and the Web 2017 Conference Cleveland, OH, 2017-04

Slides will be available at

http://www.getty.edu/research/tools/vocabularies/training.html

Microthesaurus: designated subset of a thesaurus that is capable of functioning as a complete thesaurus.

-- ISO25964-2:2013



The needs of controlled vocabularies in the Semantic Web era

Situa tions: no controlled values

→ controlled

controlled, but local → standard or popular vocabs controlled, standard vocab, but not on LOD

- Need to populate controlled vocabs in dataset
- Need to map to standard vocabs

 Need to have a LOD vocabulary (with URIs)



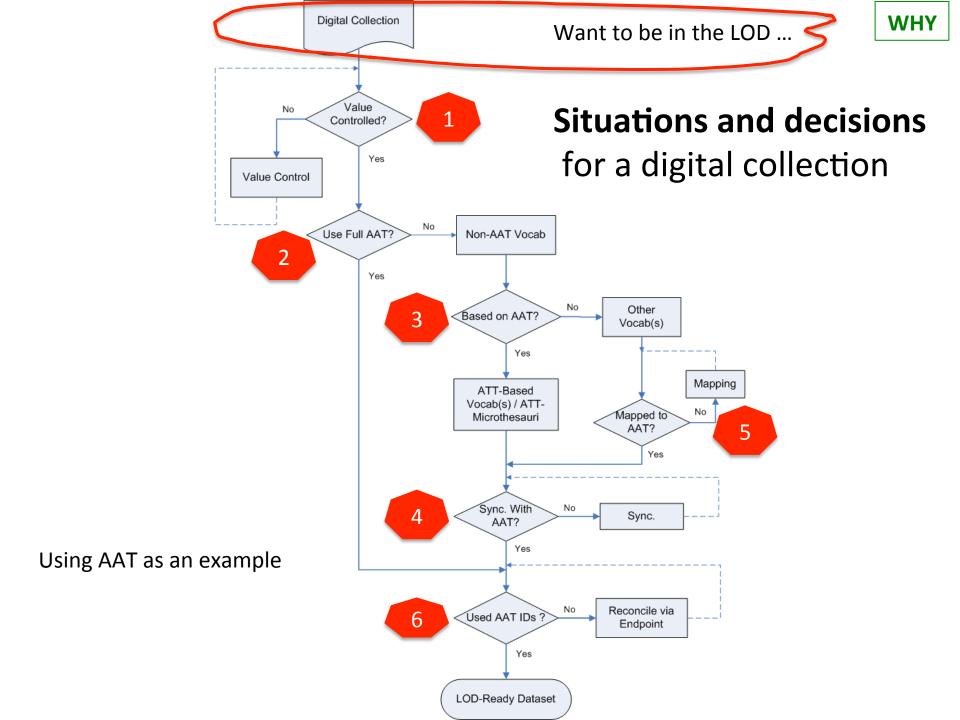
Build new vocabs based on AAT

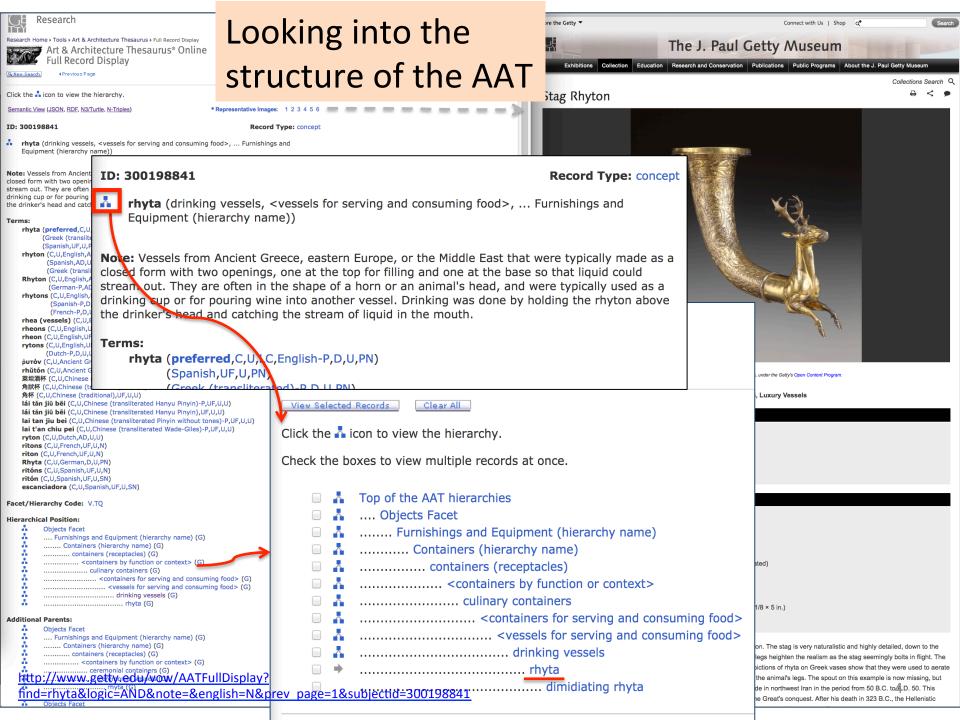


Map to selected parts of AAT



Export RDF data
using AAT LOD for
data values





<u>Facet</u>: Objects

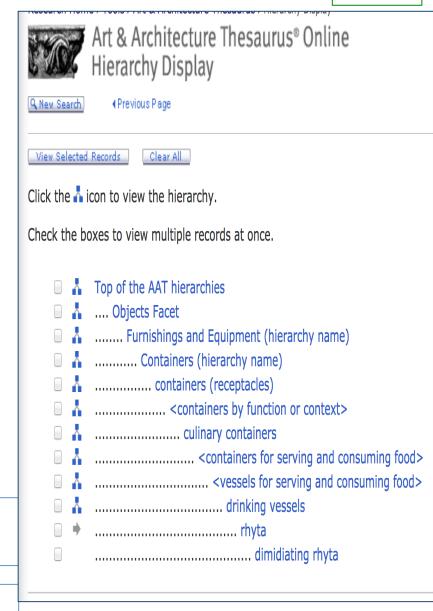
<u>Hierarchy</u>: Furnishing and Equipment

<u>Concept</u>: containers (receptacles)

<u>Guide term</u>: <containers by form>

concept:
vessels (containers)

<u>concept</u>: rhyta



Facet: Objects The units (facets, hierarchies) were recommended to be used e.g., by the *Canadian Heritage Information*Network (CHIN)

Hierarchy:

Furnishing and Equipment

Concept:

containers (receptacles)

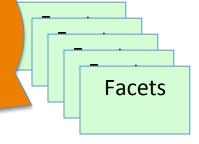
Guide term:

<containers by form>

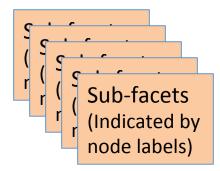
concept:

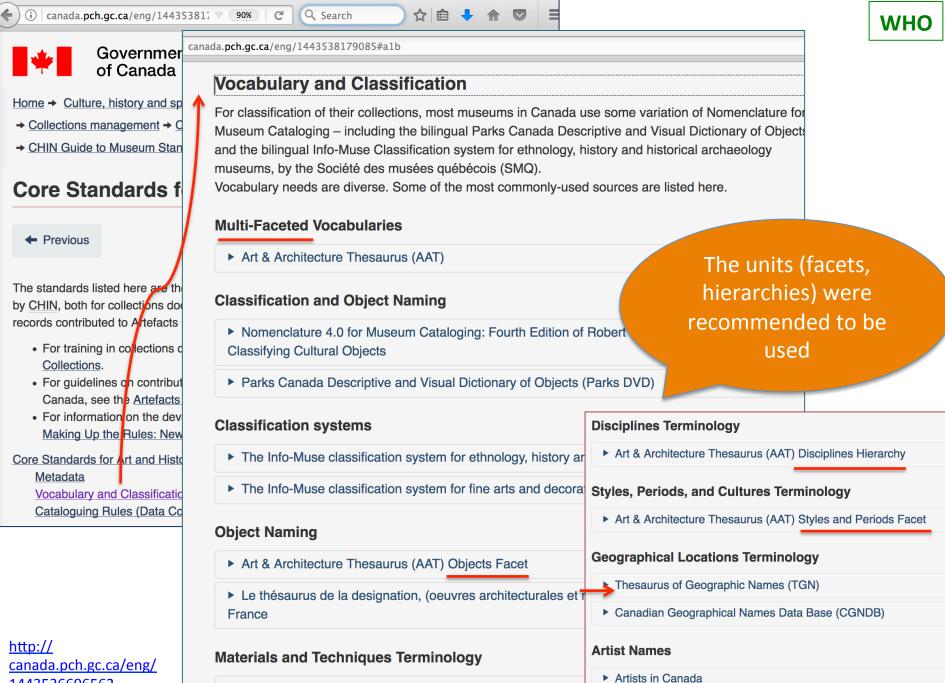
vessels (containers)

<u>concept</u>: rhyta



[large] Hierarchies (full coverage, deep layer)





Union List of Artist Names (ULAN)

► Art & Architecture Thesaurus (AAT) Materials Facet

1443536696562



Demo 1.

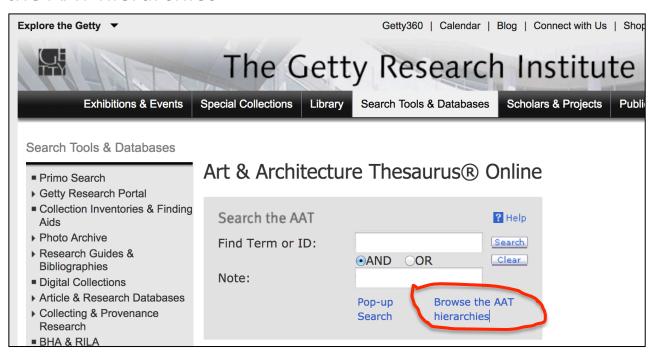
Use a <Guide Term>
to obtain
all concept URIs and
preferred terms
in the hierarchies
(for a microthesaurus
or a pick list)

Microthesaurus: designated subset of a thesaurus that is capable of functioning as a complete thesaurus.

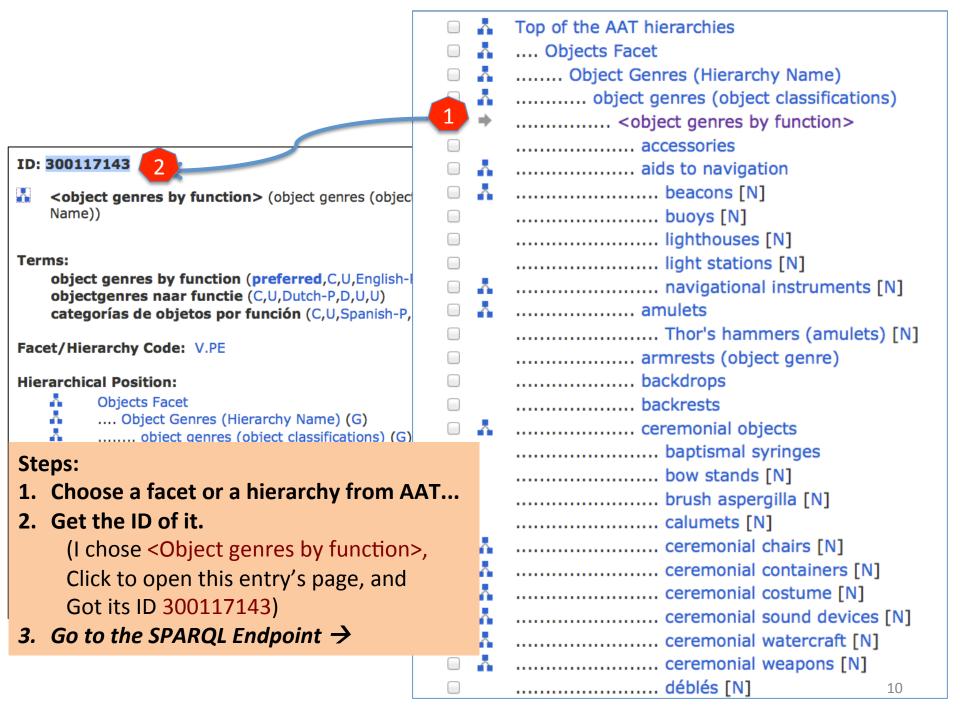
-- ISO25964-2:2013

■	Top of the AAT hierarchies		
	Objects Facet		
	Object Genres (Hierarchy Name)		
	object genres (object classific	rations)	
□ ⇒	<object by="" function<="" genres="" td=""><td>i></td><td></td></object>	i>	
	accessories	aat:300264447	agere Ifa@en
	aids to navigation	dd::000204417	agoro nason
	beacons [N]	aat:300262796	aids to navigation@en
	buoys [N]		
	lighthouses [N]	aat:300210466	aigrettes (plumes)@en
	light stations [N]	aat:300181617	airport beacons@en
	navigational instrun		alaman bassas Osa
	amulets	aat:300181651	airway beacons@en
	Thor's hammers (ar	aat:300263682	akonkromfi@en
	armrests (object genre		
	backdrops	aat:300210415	albs@en
	backrests	aat:300198819	alms dishes@en
	ceremonial objects	dd::000100010	dillo distigatori
	baptismal syringes	aat:300210416	almuces (hoods)@en
	bow stands [N]		alasta Bara
	brush aspergilla [N]	aat:300263075	aloalo@en
	calumets [N]	aat:300391092	altar bells@en
	ceremonial chairs [1		
	ceremonial containe	aat:300391083	altar candlesticks@en
	ceremonial costume	aat:300262580	altar crosses@en
	ceremonial sound d	dat.300202300	altal Closses@ell
	ceremonial watercra	aat:300198805	altar cruets@en
	ceremonial weapons déblés [N]	aat:300075940	altarpieces@en
	[]		
		aat:300264259	alusi@en
		aat:300210417	amices@en
		aat:300198899	ampullae@en
		aat:300266585	amulets@en

- 1. Go to AAT site http://www.getty.edu/research/tools/vocabularies/aat/
- 2. Choose Browse the AAT hierarchies

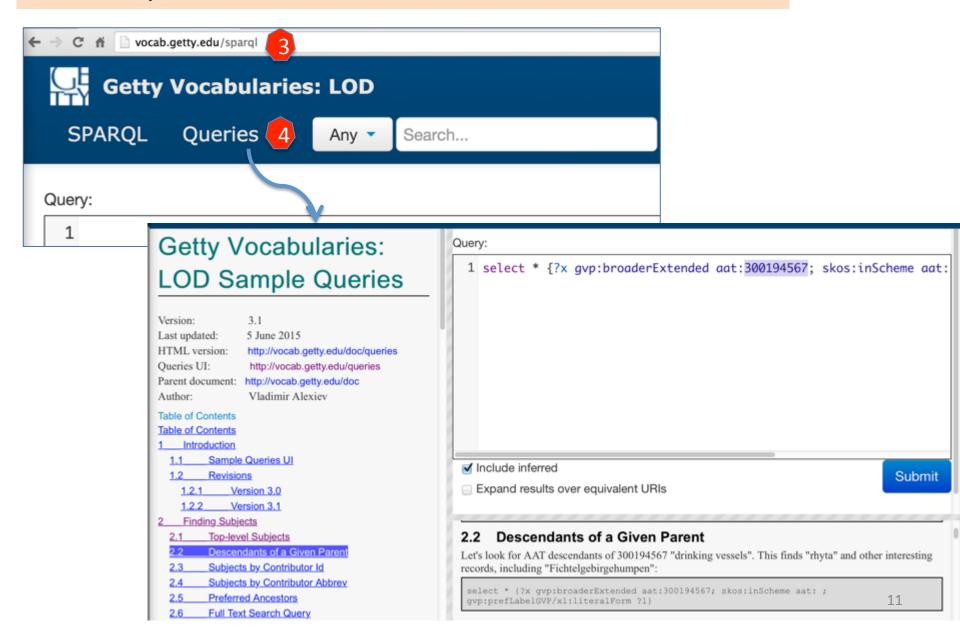


3. Explore one of the facet, all the way down, and see how far you could go.



Steps:

- 3. Go to Getty Vocab LOD SPARQL Endpoint: http://vocab.getty.edu/sparql
- 4. Choose 'Queries'.



Getty Vocabularies: LOD Sample Queries

Version: 3.1

Last updated: 5 June 2015

HTML version: http://vocab.getty.edu/doc/queries
Queries UI: http://vocab.getty.edu/queries
Parent document: http://vocab.getty.edu/doc

Author: Vladimir Alexiev

Table of Contents

Table of Contents

1 Introduction

1.1 Sample Queries UI

1.2 Revisions

1.2.1 Version 3.0

1.2.2 Version 3.1

Prinding Subjects

2.1 Top level Subjects

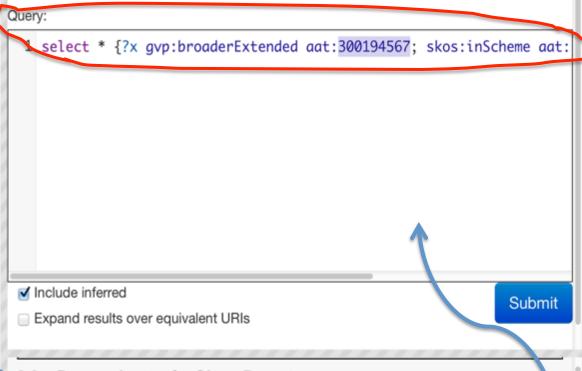
2.2 Descendants of a Given Parent

2.3 Subjects by Contributor Id

2.4 Subjects by Contributor Abbrev

2.5 Preferred Ancestors

2.6 Full Text Search Query



2.2 Descendants of a Given Parent

Let's look for AAT descendants of 300194567 "drinking vessels". This finds "rhyta" and other interesting records, including "Fichtelgebirgehumpen":

select * {?x gvp:broaderExtended aat:300194567; skos:inScheme aat: ;
gvp:prefLabelGVP/x1:literalForm ?1}

Steps:

5

- 5. Choose "Descendants of a Given Parent" from the template, click.
- → Now, the template's text will show on the right.
- 6. Click 'SPARQL" to get the query text up.

SPARQL

Steps

7. Replace the ID 300117143 (= <Object genres by function>) in the Query template. [you may modify to add more requests if you know how.]

```
Query:

1 select * {?x gvp:broaderExtended aat:300194567; skos:inScheme aat: ; gvp:prefLabel

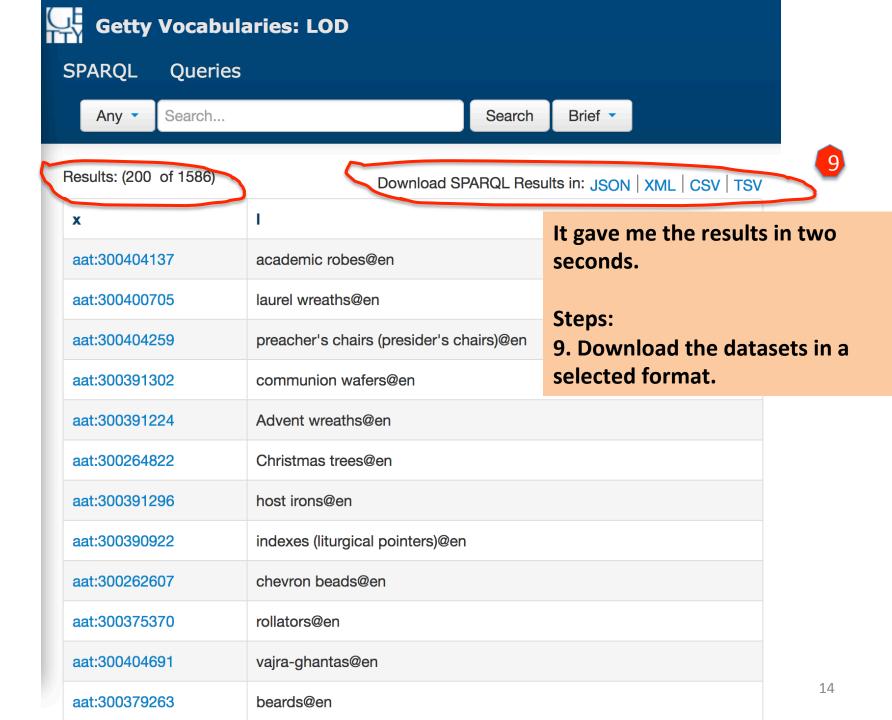
Query:
```

1 select * {?x gvp:broaderExtended aat:300117143; skos:inScheme aat: ; gvp:prefLabel

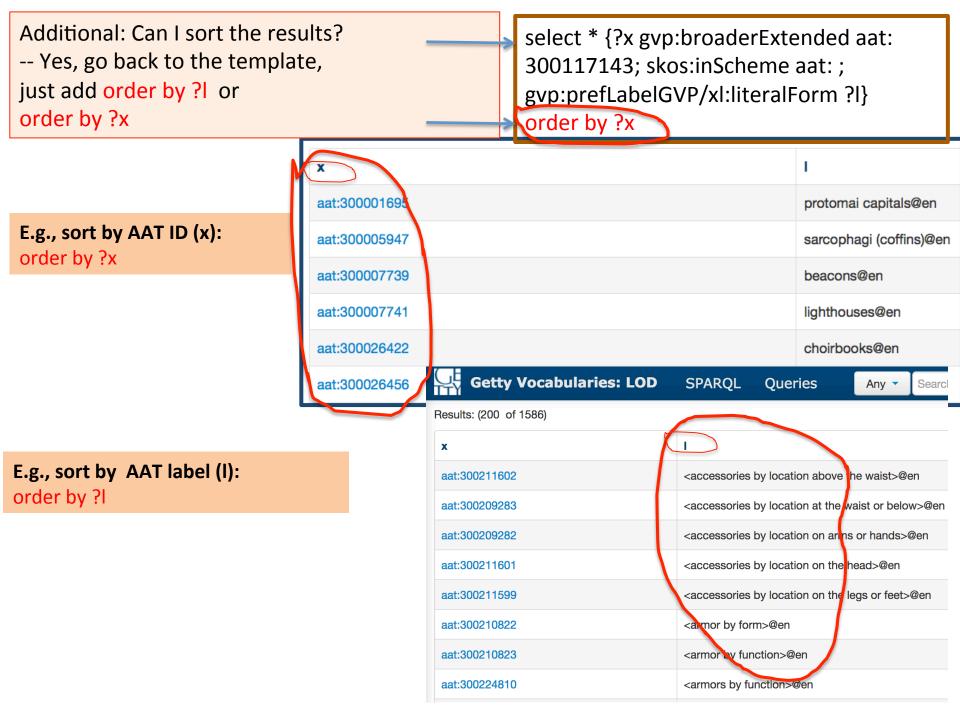
- ✓ Include inferred
- Expand results over equivalent URIs







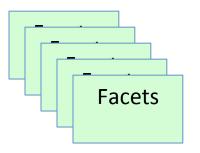
	Top of the AAT hierarchies				What did I get?
	Objects Facet				
	Object Genres (Hierarchy Name)				The URIs for every concept
	object genres (object classifications)		D4 D 0		in this <object by<="" genres="" th=""></object>
□ →	<object by="" function="" genres=""></object>	S	PARQ	L Query	function> family.
	accessories				
	aids to navigation		Results fo	or # 4.1.2 Desce	ndants (100 of 523)
	beacons [N]				
	buoys [N]				Develop I OPAROL Provide in
	lighthouses [N]				Download SPARQL Results in: JSON XML
	light stations [N]	Х			1
	navigational instruments [N]	a	at:300161	886	striking blocks@en
	amulets	a	at:300202	542	keepsakes (books)@en
	Thor's hammers (amulets) [N]	a	at:300220	519	religious texts@en
	armrests (object genre)	a	at:300210	422	buskins (stockings)@en
	backdrops		-4-000404	000	
	backrests	a	at:300181	600	aeronautical beacons@en
	ceremonial objects	a	at:300007	739	beacons@en
	hantismal syringes	ctio	182	\$41	lighthouse lamps@en
	• beacons	00.0.	3 17	741	lighthouses@en
	aeronautical beaco	ns	180	588	Hanukkah lamps@en
	airport beacons		190	801	votive lamps@en
	airway beacons				
	landmark beaco				
	obstruction bea	cons	5		
	lighthouse lamps				
	lightships [N]				
U.	L-3				



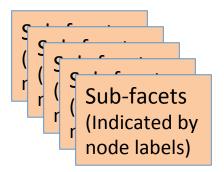
HOW

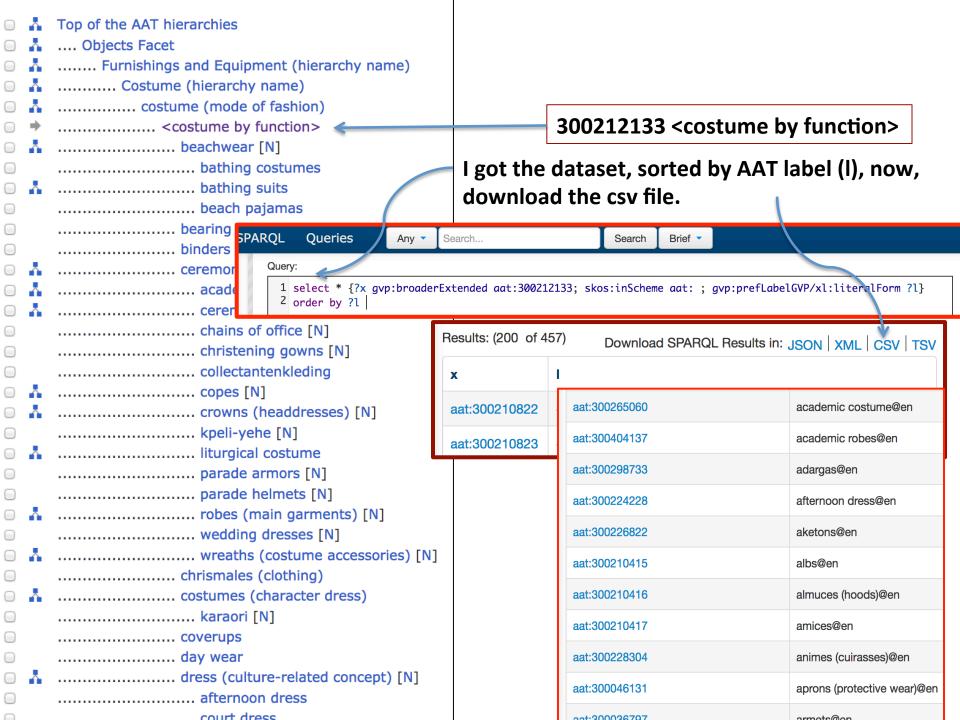
Demo 2.
Download the datasets for

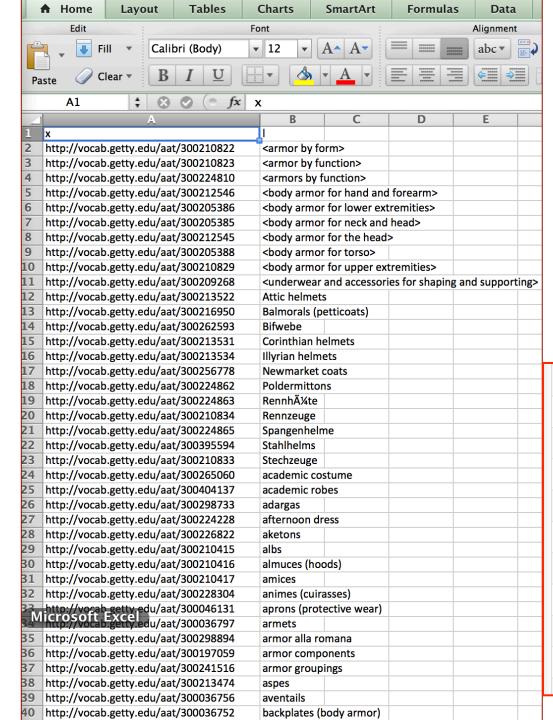
300212133 <costume by function>



[large] Hierarchies (full coverage, deep layer)







Open the CSV file from spreadsheet on my laptop

300212133 <costume by function>

academic costume@en

afternoon dress@en

almuces (hoods)@en

animes (cuirasses)@en

aprons (protective wear)@en

aketons@en

albs@en

amices@en

aat:300404137	academic robes@en
aat:300298733	adargas@en

aat:300265060

aat:300224228

aat:300226822

aat:300210415

aat:300210416

aat:300210417

aat:300228304

aat:300046131

+-200026707

academic robes

Source:http://vocab.getty.edu/aat/300404137

gvp:broaderExtended

dcterms:license

prov:wasGeneratedBv

Subject (83) **Predicate** Object Download in: JSON JSONLD RDF N3/Turtle N-Triples Website Hierarchy Inference Implicit only Statements in which the resource exists as a subject. **Predicate Object** gvp:Subject, skos:Concept rdf:type academic robe@en, academic robe@en, academische toga@nl, academische toga's@nl rdfs:label aat:300265060, aat:300404136 gvp:broader aat:300117143, aat:300178802, aat:300185711, aat:300185712, aat:300209261, aat:300209263, gvp:broaderGenericExtended aat:300209852, aat:300210387, aat:300212133, aat:300234117, aat:300264092, aat:300264551,

aat:300117143, aat:300178802, aat:300183711, aat:300185712, aat:300209261, aat:300209263,

aat:300209852, aat:300210387, aat:300212133, aat:300234117, aat:300264092, aat:300264551,

aat:300265060, aat:300266639, aat:300404136

aat:300265060, aat:300266639, aat:300404136

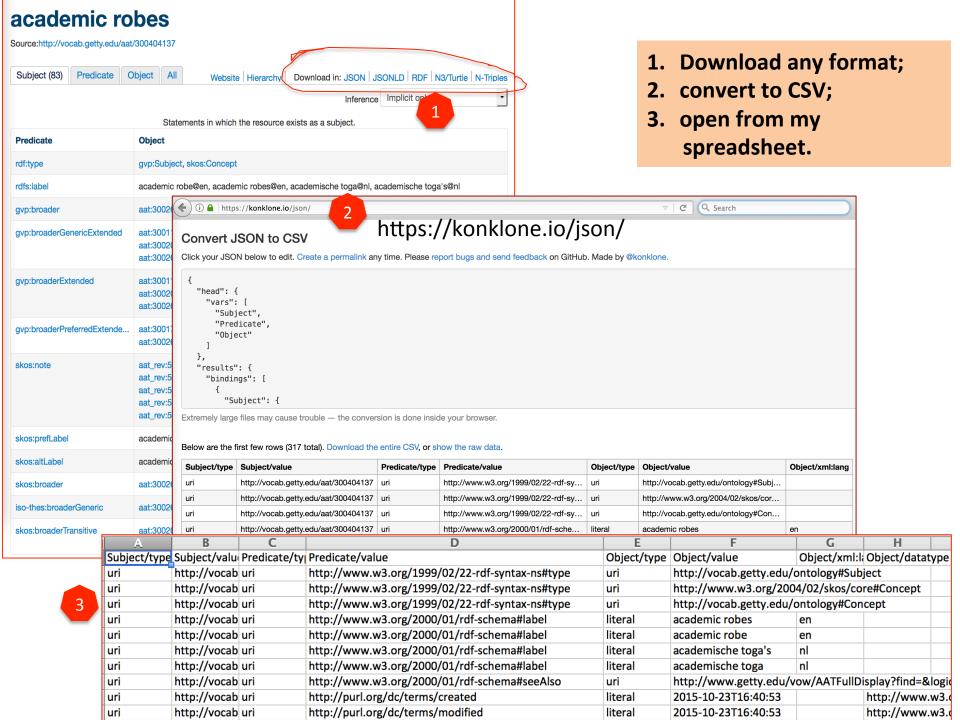
http://opendatacommons.org/licenses/by/1.0/

aat_rev:5004034709

If I want to get data for each concept in this family ...

300212133 <costume by function>

gvp:broaderPreferredExtende aat:300178802, aat:300209261, aat:300209263, aat:300209852, aat:300264092, aa		aat:300265060	academic costume@en
	aat:300266639, aat:300404136		uoudonno ocolumoeon
skos:note	aat_rev:5004034708, aat_rev:5004034709, aat_rev:5004034712, aat_rev:50040347 aat_rev:5004034715, aat_rev:5004034717, aat_rev:5004034718, aat_rev:50040347 aat_rev:5004034720, aat_rev:5004044324, aat_rev:5004044325, aat_rev:5004044324 aat_rev:5004044582, aat_rev:5004055794, aat_rev:5004055795, aat_rev:5004055794	aat:300404137	academic robes@en
		aat:300298733	adargas@en
		aat:300224228	afternoon dress@en
skos:prefLabel	academic robes@en, academische toga's@nl	aat:300226822	aketons@en
skos:altLabel	academic robe@en, academische toga@nl	aat:300210415	albs@en
skos:broader	aat:300265060, aat:300404136	uanissis 1.15	
iso-thes:broaderGeneric	aat:300265060, aat:300404136	aat:300210416	almuces (hoods)@en
skos:broaderTransitive	aat:300209263, aat:300404136, aat:300265060, aat:300178802, aat:300185712, aa aat:300266639, aat:300234117, aat:300209852	aat:300210417	amices@en
		aat:300228304	animes (cuirasses)@en
cc:license http://ope	endatacommons.org/licenses/by/1.0/		
void:inDataset http://voo	p://vocab.getty.edu/dataset/aat	aat:300046131	aprons (protective wear)@en



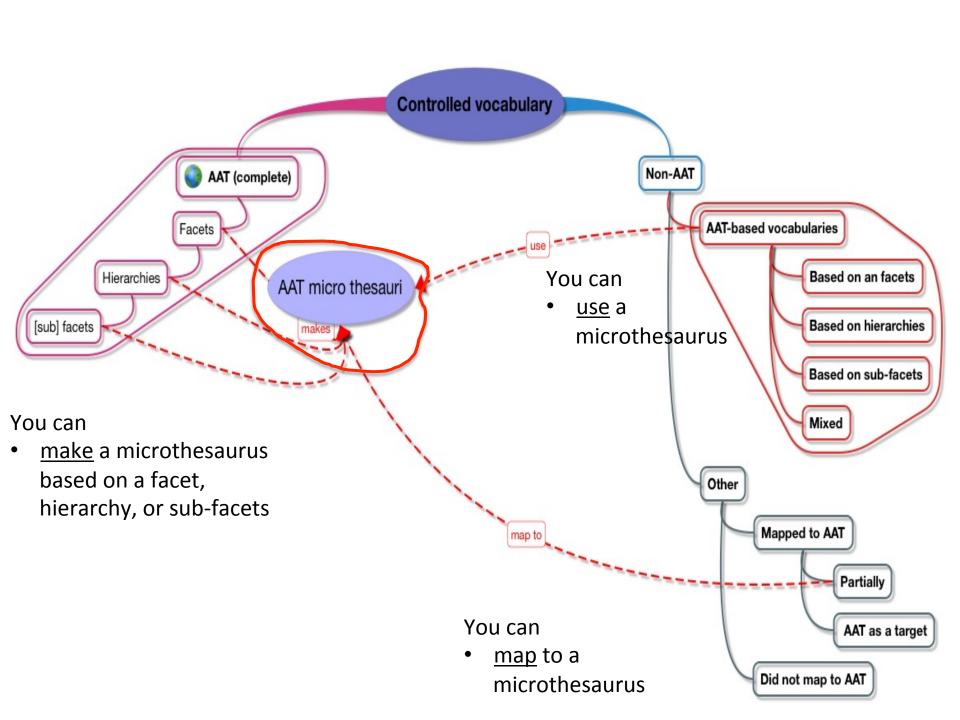
Miscrothesauri -- Summary

- 1. Choose the facet or hierarchy you like to start;
- 2. Find the ID of that concept.
- 3. Use this template to get the URIs and labels:

```
select * {?x gvp:broaderExtended aat: 300194567; skos:inScheme aat: ; gvp:prefLabelGVP/xl:literalForm ?l} order by ?x
```

- Replace the ID in the Query template
- Submit
- Get the URIs and labels in under this guide term.
- Sort by order (column x)

- 4. Use a tool that can treat CSV (or convert from JSON to CSV) to view and manage.
- 5. Extend from these URIs to get dataset for any concept.
- 6. Use other queries to get different datasets.



Could a LOD KOS dataset be considered

- as a knowledge base?
- as the foundation of a network analysis?
- as the building blocks of a framework



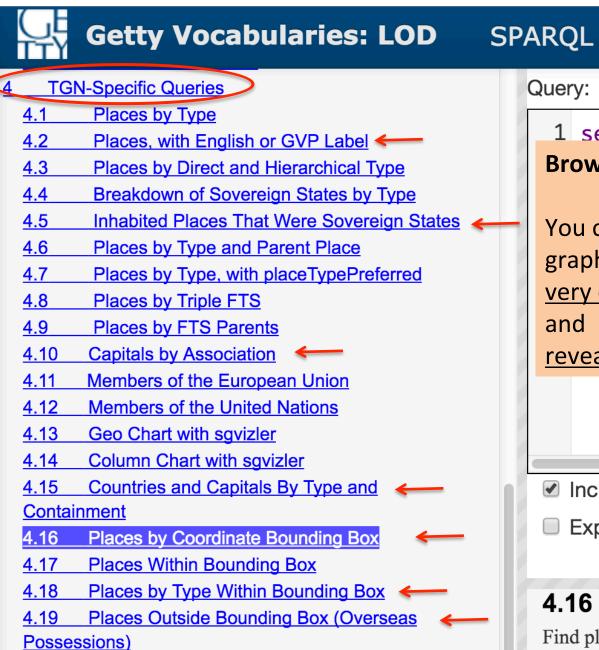
For Researchers

Demo 3. Using TGN

Getty Thesaurus of Geographic Names

- 1. Go to http://vocab.getty.edu/queries
 - Refresh the query website
- 2. Find section 4. TGN-specific queries

Demo: Look for castles around The Netherlands (within the boundary of 50.787185 3.389722 53.542265 7.169019)



Query:

1 select ?place

Browse the examples of queries

Queries

You can obtaining special RDF graphs or datasets for very complicated questions, and revealing unknown relationships

- Include inferred
- Expand results ove

4.16 Places by

Find places whose coordi

select ?place ?name

Places Nearby Each Other

4.20



- (1) go to 4.18,
- (2) click on that SPARQL sign for 4.18,
- (3) submit. Note: Since this is a complicated

query, it will run a few seconds.

Query: Look for castles around The Netherlands (within the boundary of 50.787185 3.389722 53.542265 7.169019)

i

 ✓ vocab.getty.edu

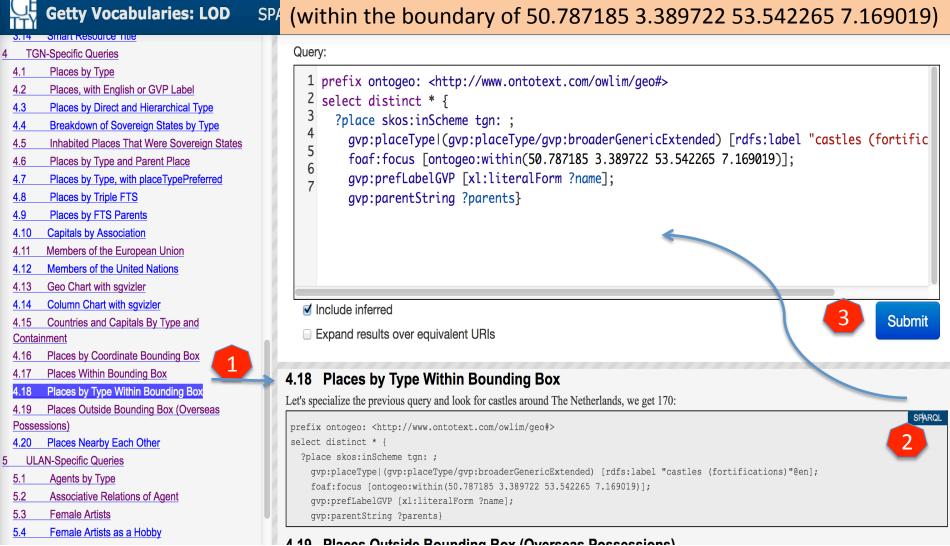
Getty Vocabularies: LOD

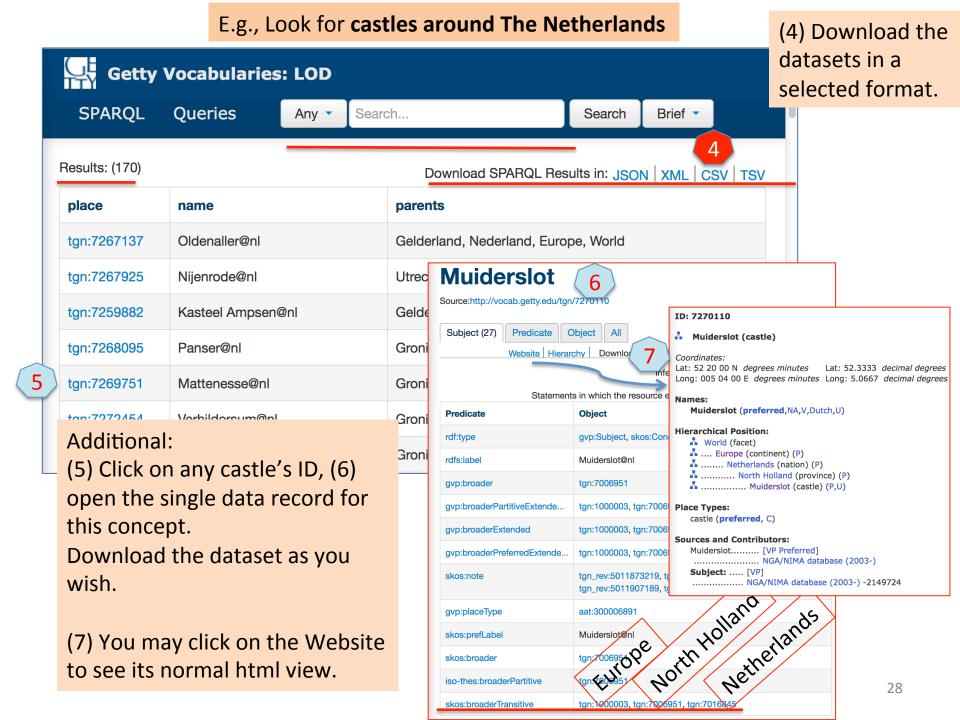
The Getty Vocabularies

Welcome to the SPARQL endpoint vocab.getty.edu

SPAROL

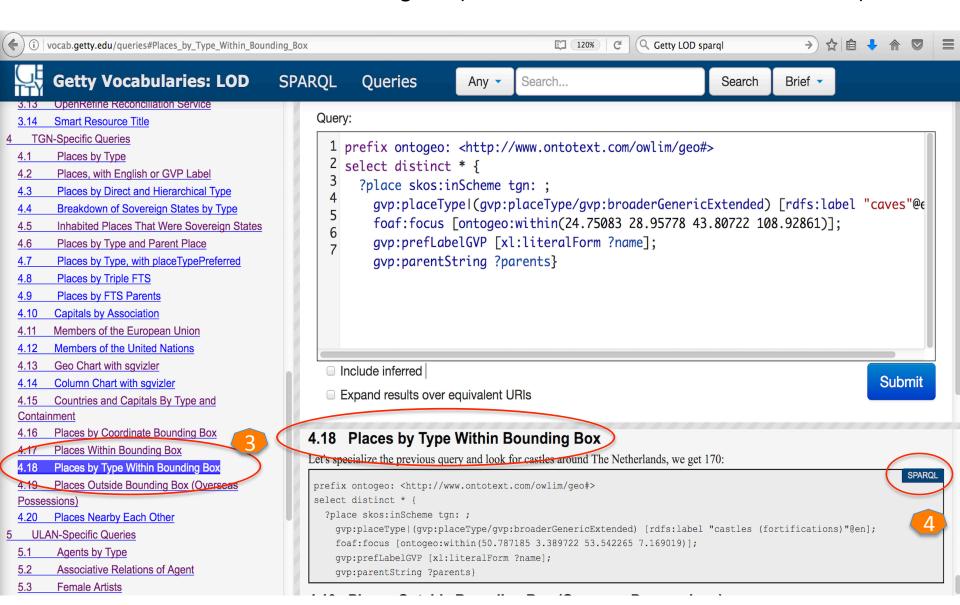
Oueries





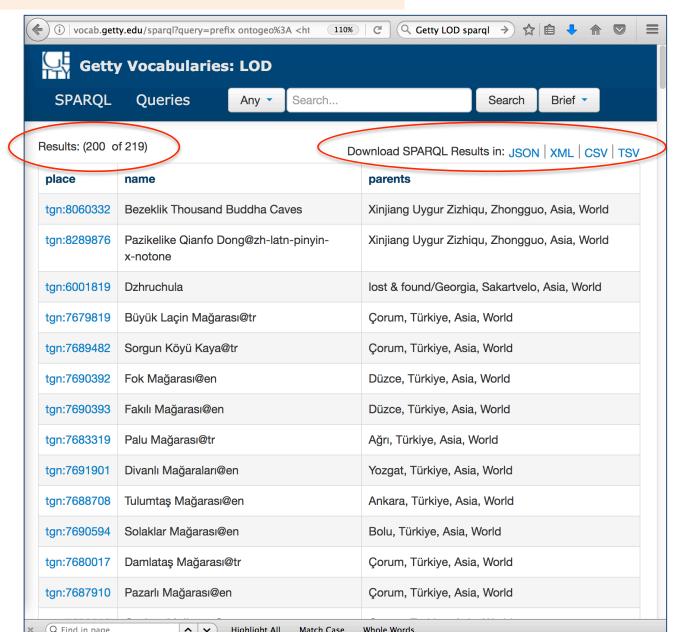
Another example: Look for caves on or around the Silk Road

caves within bounding box (24.75083 28.95778 43.80722 108.92861)

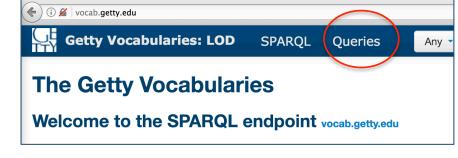


Query a specific place type (e.g., caves) in a geographic boundary

Got the results & downloadable datasets:



caves within(24.750 83 28.95778 43.80722 108.92861)



Demo 4. Using ULAN

Union List of Artist Names

- 1. You may go to ULAN to find the ID of someone you want to search for.

 http://www.getty.edu/research/tools/vocabularies/ulan/
- 2. We will be at the same query templates page http://vocab.getty.edu/queries.
 - Refresh the query website
- 3. Find section 5. ULAN-specific queries

Demo: 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.

Explore the Getty ▼ Getty360 | Calendar | Blog | Connect with Us | Shop

Library

Special Collections

The Getty Research Institute

Search Tools & Databases

Scholars & Projects Publication

http://www.getty.edu/research/tools/vocabularies/ulan

Search Tools & Databases

Exhibitions & Events

■ Primo Search

- ▶ Getty Research Portal
- Collection Inventories & Finding Aids
- ▶ Photo Archive
- Research Guides & Bibliographies
- Digital Collections
- ▶ Article & Research Databases
- ▶ Collecting & Provenance Research
- BHA & RILA
- → Getty Vocabularies

Union List of Artist Names® Online

Search the ULAN			? Help
Find Name or I. 1	Frank Lloyd	Wright	Search
Role: Lookup Nationality: Lookup			Clear
	Pop-up Search	Browse th	

Page Click the \(\frac{1}{4} \) icon to view the hierarchy. Check boxes to view multiple records at once. ■ ♣ Frank Lloyd Wright Building Conservancy (American organization, established 1990) [500231702] (Corpora ■ Frank Lloyd Wright Foundation (American foundation, founded 1940) [500243110] (Corporate B Frank Lloyd Wright Memorial Foundation ■ ♣ Frank Lloyd Wright Home and Studio Foundation (American repository, Oak Park, active 1974-2000) [500303810] Frank Lloyd Wright Preservation Trust (American society, contemporary) [500271821] (Corporate Bodie ■ ♣ Frank Lloyd Wright School of Architecture (American architecture school, contemporary) [500249649] (Con ■ National Foundation for the Study of Frank Lloyd Wright (American repository, Detroit, contemporary) [500306742] (Corp Wright, Frank Lloyd (American architect, 1867-1959) [500020307] (Persons, Artists) Frank Lloyd Wright 🔲 🔥 Wright, Lloyd (American architect and scenographer, 1890-1978) [500009812] Wright, Frank Lloyd

From ULAN, find the ID of someone you want to search for.



Research

Research Home ▶ Tools ▶ Union List of Artist Names ▶ Full Record Display



Union List of Artist Names® Online Full Record Display

Q New Search

◆ Previous Page

Click the $\frac{1}{4}$ icon to view the hierarchy.

Semantic View (JSON, JSONLD, RDF, N3/Turtle, N-Triples)

ID: 500020307

3

Wright, Frank Lloyd (American architect, 1867-1959)

Note: Architect, draftsman, furniture and interior designer ar ecognized as one of the most important figures in 20th centuarchitects whose success and reputation are measured in par everyday landscape, Wright remained an individualist who most 20th century architecture. Instead, his fame is derived from within an environment and within space itself. He is best known office buildings, and innovative furniture and chinaware designate.

Names:

Wright, Frank Lloyd (preferred, V, index, LC)

Frank Lloyd Wright (V,display)

Frank Lloyd Lincoln Wright (V)

Rait, Frank Lloid (V)
Rayt, Frank Luyd (V)

Nationalities:

American (preferred)

undetermined



Getty Vocabularies: LOD

SF

4.20 Flaces Nearby Lacit Other

5	ULAN-S	pecific C	ueries

- 5.1 Agents by Type
- 5.2 Associative Relations of Agent
- 5.3 Female Artists
- 5.4 Female Artists as a Hobby
- 5.5 Native American Painters
- 5.6 Names of Native American Painters
- 5.7 Architects Born in the 14th or 15th Century
- 5.8 Indian and Pakistani Architectural Groups
- 5.9 Non-Italians Who Worked in Italy
- 5.10 Artists Associated to a Given Patron or His Family
- 5.11 German, Dutch, Flemish printmakers, listed with their teachers
- 5.12 Artists Whose Identity May be Associated or Confused With Another
- 5.13 Ordered Hierarchy of Given Subject
- 5.14 Ancient Artists or Groups by Nationality
- 5.15 Art Repositories in the USA by State
- 5.16 Popes and Their Reigns
- 5.17 Pope Reign Durations
- 5.18 Life Events

Go to the same query templates page

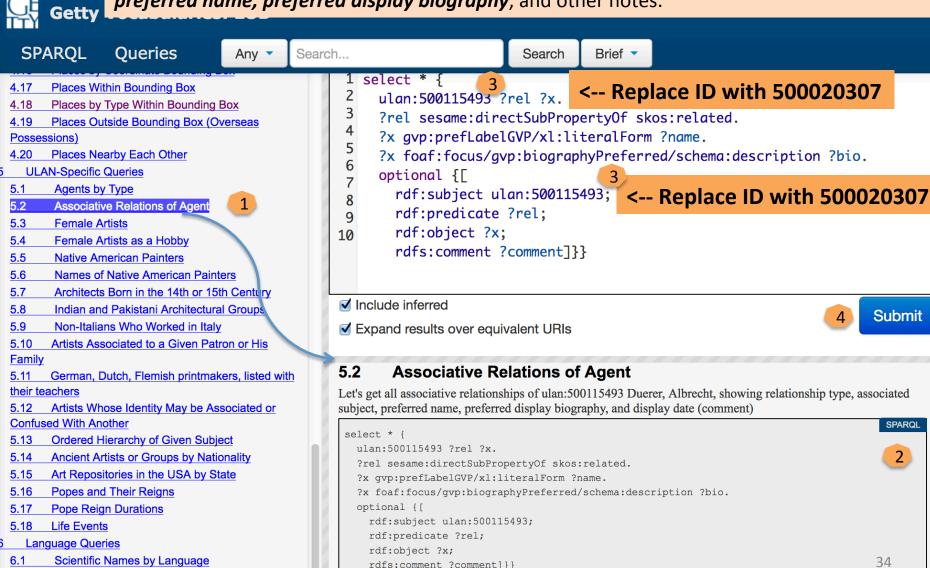
http://vocab.getty.edu/queries

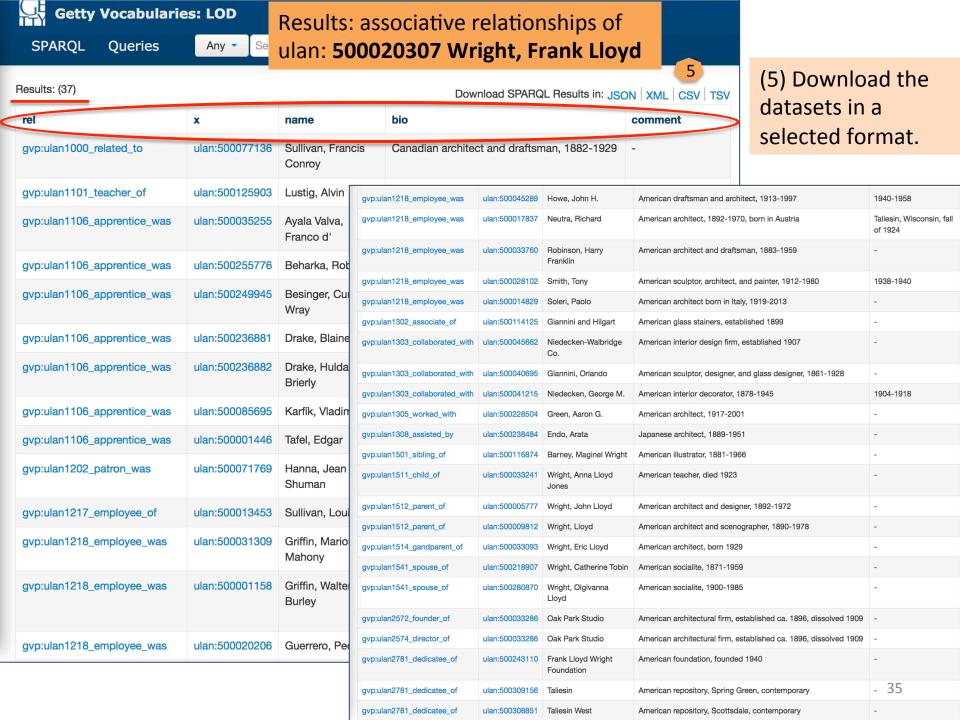
Find the section for ULAN.

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.





Summary: Follow the demos to obtain your datasets

- Demo 1. Use a <Guide Term> to obtain all concept URIs and preferred terms in the hierarchies (for a microthesaurus or a pick list), e.g., for <Object genres by function> (AAT ID 300117143)
- Demo 2. Download the datasets for <costume by function> (AAT ID 300212133)
- Demo 3. Use TGN to obtaining special RDF graphs or datasets for <u>very</u> <u>complicated questions</u>, e.g., castles around The Netherlands (within the boundary of 50.787185 3.389722 53.542265 7.169019)
- **Demo 4.** Use ULAN to 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.



How to learn more about Linked Data?

Go to: http://explore.dublincore.net/



LD4PE Products

1. Competency Index

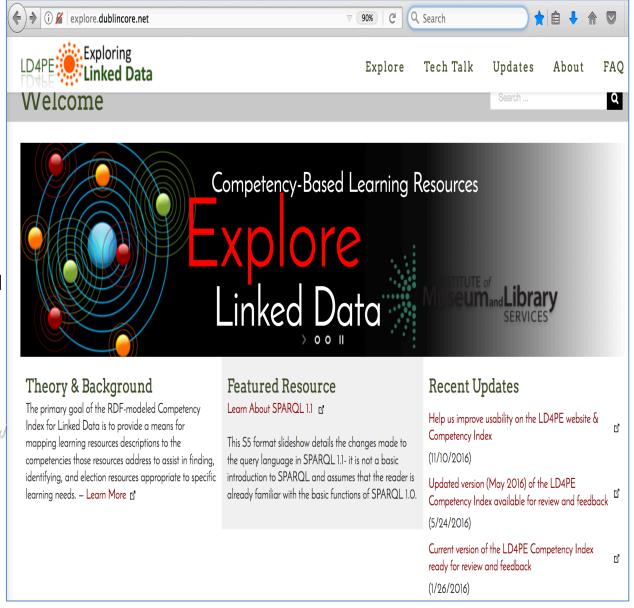
- Four-levels of granularity (Topic Cluster, Topic, Competency, Benchmark)
- Modeled after Achievement Standards Network (ASN)

2. Resource Description Tool

- Resources are Indexed at the Topic and Competency Level
- Saved to triple store connected to Word Press site

3. Exploring Linked Data– a Website

- Browse Competency index
- Search for Resources
 - By Competency
 - By Keyword
- Saved Sets of Resources
- Learning Trajectory Maps



http://explore.dublincore.net/

Create Microthesauri and other datasets from the Getty LOD Vocabularies

Marcia Lei Zeng mzeng@kent.edu

Museums and the Web 2017 Conference Cleveland, OH, 2017-04

Question?