The Role of LOD Datasets

The Getty Vocabularies and the Significance of Five-Star LOD Datasets

Marcia Lei Zeng,
Kent State University, USA

International Terminology Working Group
Getty Research Institute, L.A.
August 22 – 24, 2016

Five-Star Data ★★★★★

Sir Tim Berners-Lee, the inventor of the WWW and the initiator of Linked Data, presented a Star Scheme for measuring the rank of a dataset

★ Available on the web (whatever format) *but with an open licence, to be Open Data*

★★ Available as machine-readable structured data (e.g. excel instead of image scan of a table)

★★★ as (2) plus non-proprietary format (e.g. CSV instead of excel)

★★★★ All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff

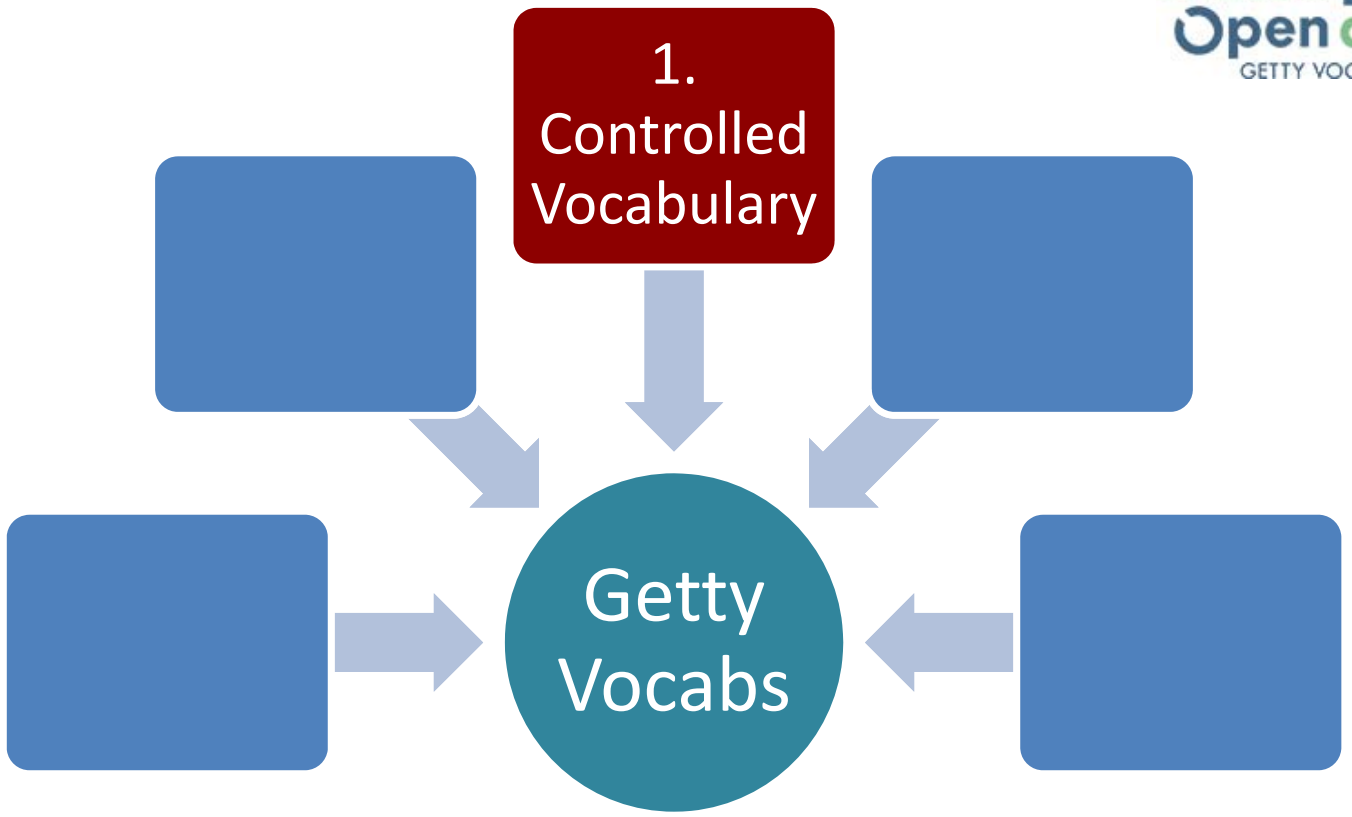
★★★★★ All the above, plus: Link your data to other people's data to provide context

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

2

What is the "Getty Vocabularies"?

(i.e., Why does any dataset need to care about it?)



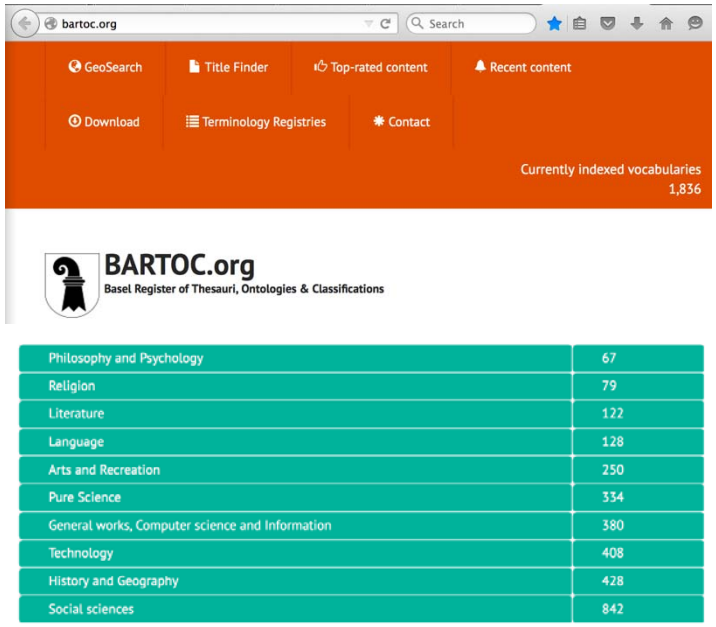
"Why Choose the Getty Vocabularies? There are so many…"

In the BARTOC registry

(thesaurus, ontology, classification)

KOS registered: **1836**

2016.05.27



http://bartoc.org/

in the Datahub

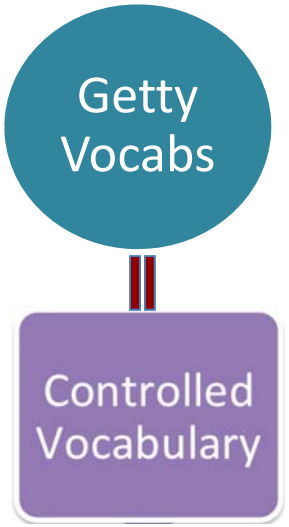
LOD KOS registered: **1251**

(about a half are ontologies)

2016.03.15

https://datahub.io/

To be a five-star LOD dataset,
one has to be already a five-star product



The Getty Vocabularies -- is a five-star vocabulary

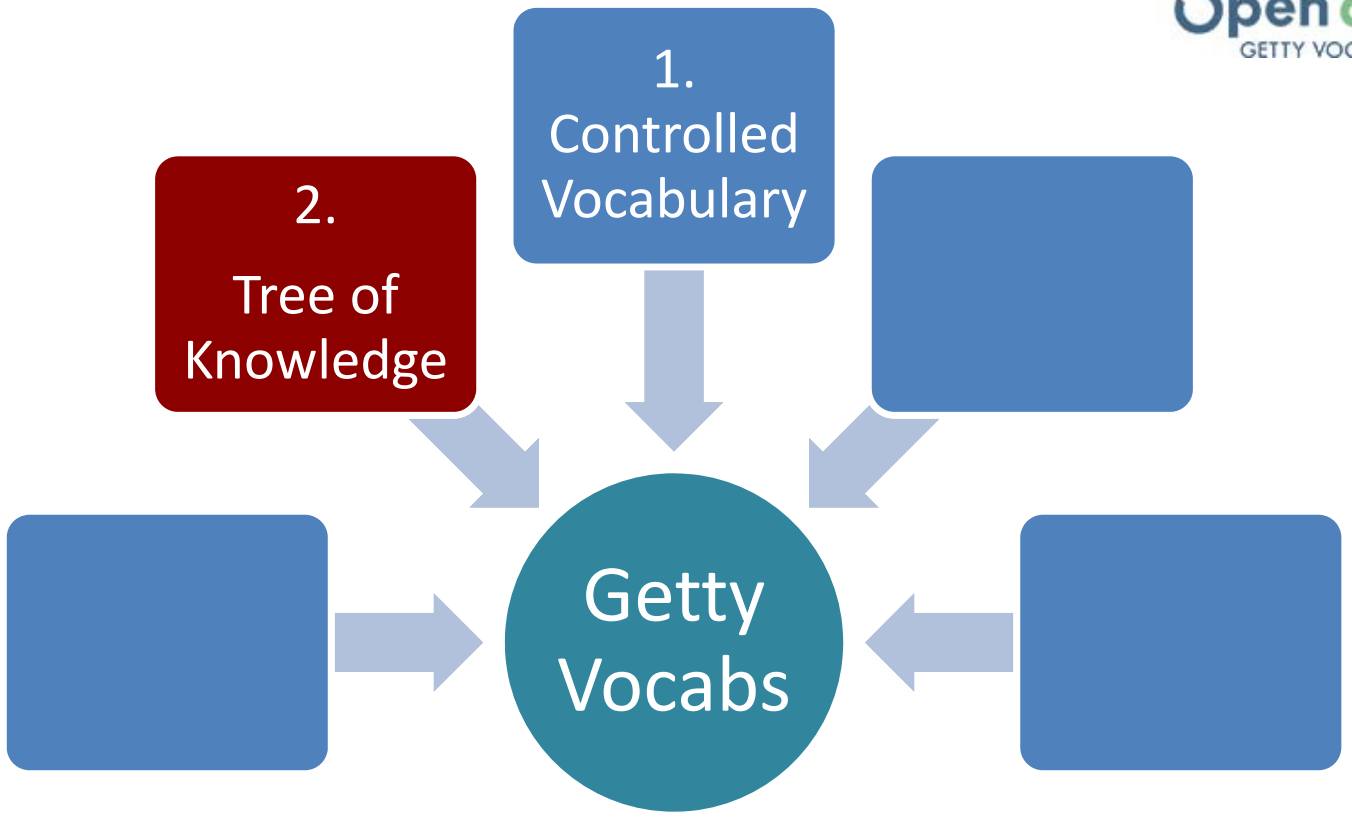
- High quality authority control of appellations representing things;
- Multilingual and multi-cultural; historical and contemporary;
- High specificity while comprehensive; continual and open-ended;
- One of the few selected vocabularies that are being:
  - recommended or required by many important metadata standards (e.g., DC., VRA Core, CCO, etc.)
  - used as examples at national and international standards for structured vocabularies (e.g., ISO25964-1 and ISO25964-2, NISO Z39.19)
  - adopted by cross-country and cross-domain data services, in addition to many institutions' (e.g., Europeana, DPLA (Digital Public Library of America))
  - widely studied by researchers. Google Scholar shows results when searching (exact match):
    - 2,110 entries for "Art and Architecture Thesaurus"
    - 3,570 for "Thesaurus of Geographic Names"
    - 89 for "Cultural Objects Name Authority"
    - 72 for "Union List of Artist Names"
    - 355 for "Getty Vocabularies" ... ...

    2016.07.20

    In comparison:
    - "Eurovoc": 2,220
    - "Library of Congress Name Authority": 768
  - ...

What is the "Getty Vocabularies"?

(i.e., Why does any dataset need to care about it?)



Porphyrian tree

Porphyry (234-ca. 305 CE)
Greek philosopher

In his *Isagoge* ("Introduction" to Aristotle's "Categories"), he

- reframed Aristotle's original predicable into a decisive list of five classes
  - genus (genos),
  - species (eidos),
  - difference (diaphoro),
  - property (idion), and
  - accident (sumbebekos).
- introduced a hierarchical, finite structure of classification

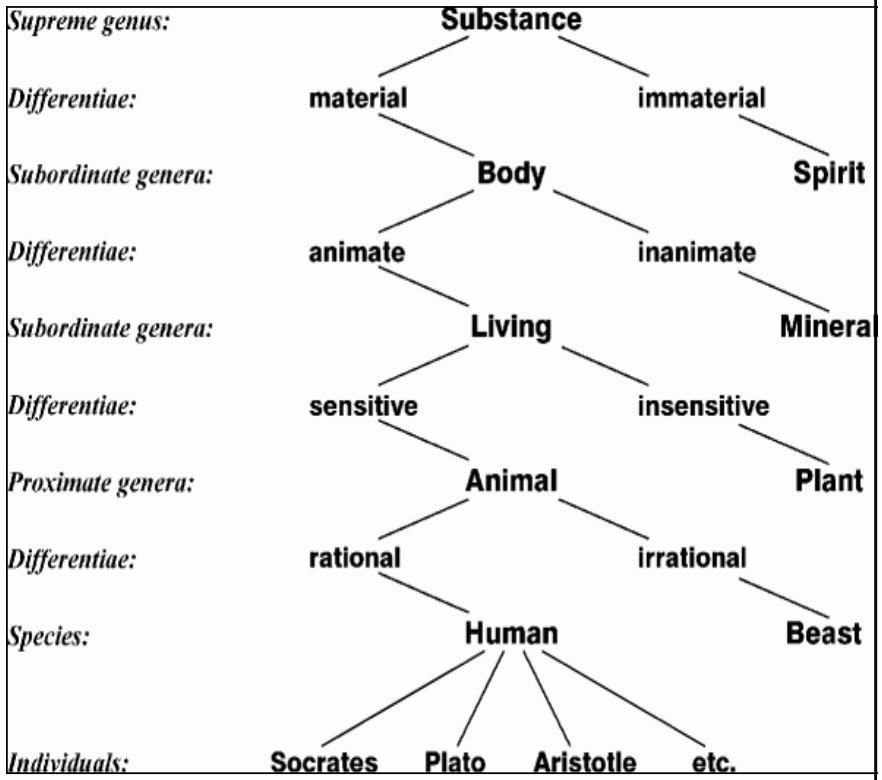
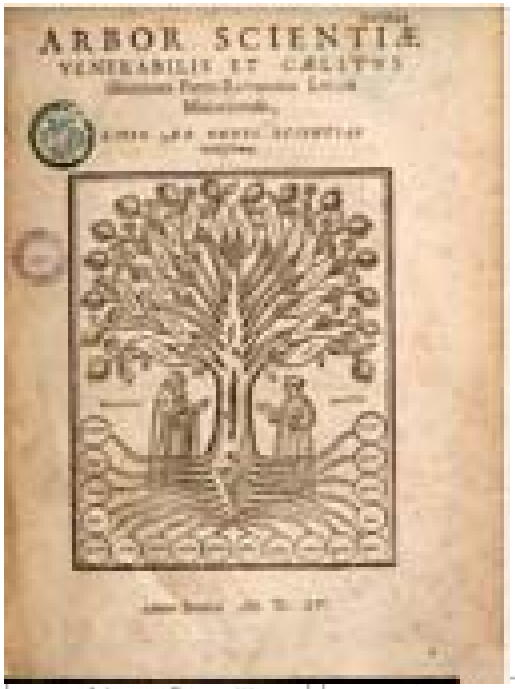


Image: A Porphyrian tree, originally draw by the 13th century logician Peter of Spain.

http://www.tertullian.org/fathers/porphyry_isagogue_01_intro.htm

https://en.wikipedia.org/wiki/Porphyrian_tree



Llull: **Tree of science**

Ramon Llull (Catalan, 1232–1315)

http://www.historyofinformation.com/expanded.php?id=3862

1295 – 1296， Ramon Llull published ***Arbor scientiae*** **(Tree of science)**

This encyclopedia and pioneering work in knowledge representation included sixteen trees of scientific domains following the initial tree called the *arbor scientiae.*

Page 123

Page 92

Page 453

Page 339

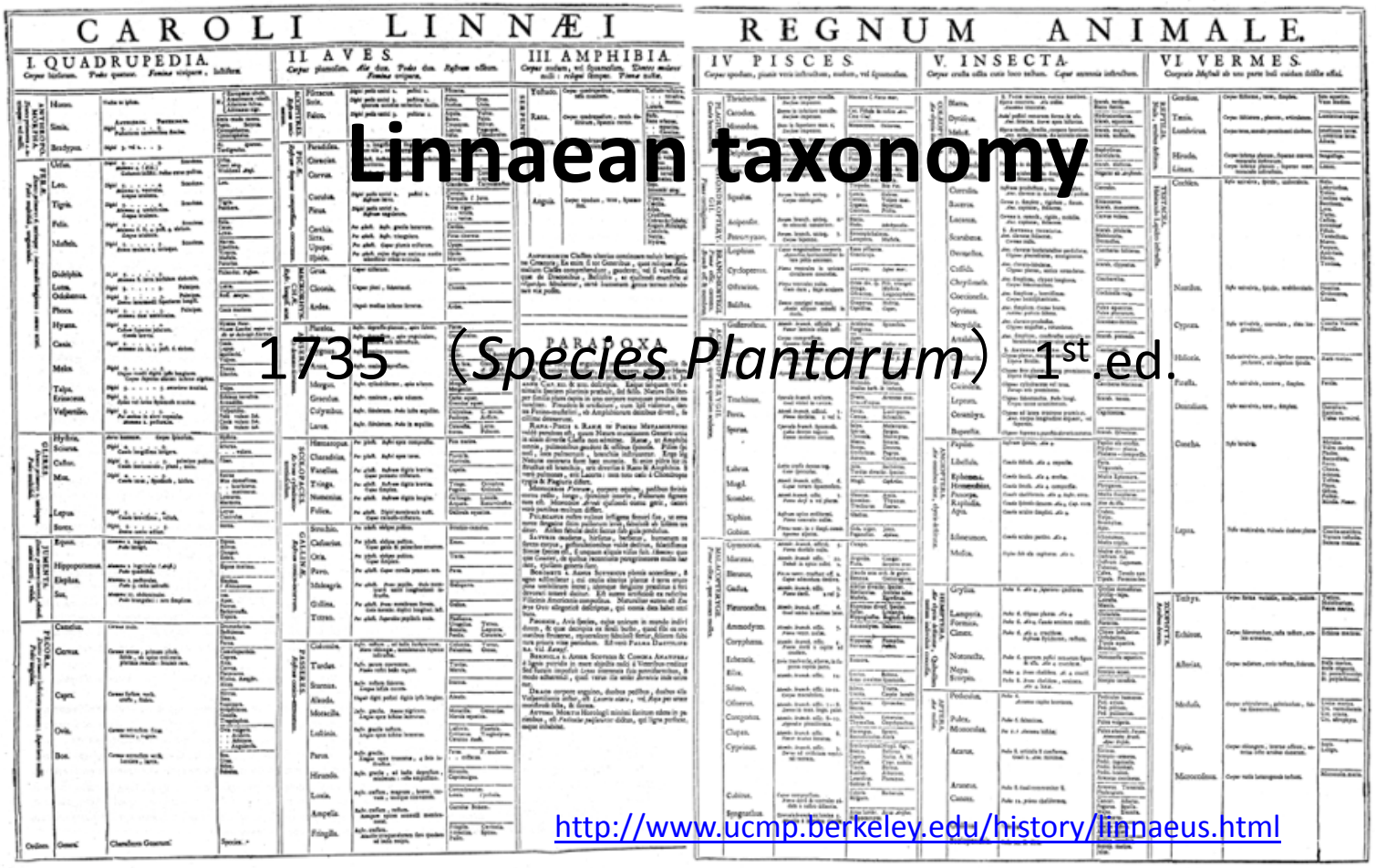
Page 181

Image source: a version published in Lyon, 1635, available through Google Books.
https://books.google.com.tw/booksid=I64oL87aiS0C&source=gbs_navlinks_s

8

Carl von Linné (1707 –1778)
(=Carolus Linnaeus)

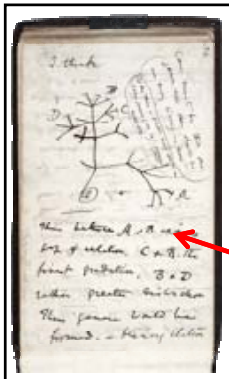
Table of the Animal Kingdom (Regnum Animale) from the 1st edition of Systema Naturæ (1735)



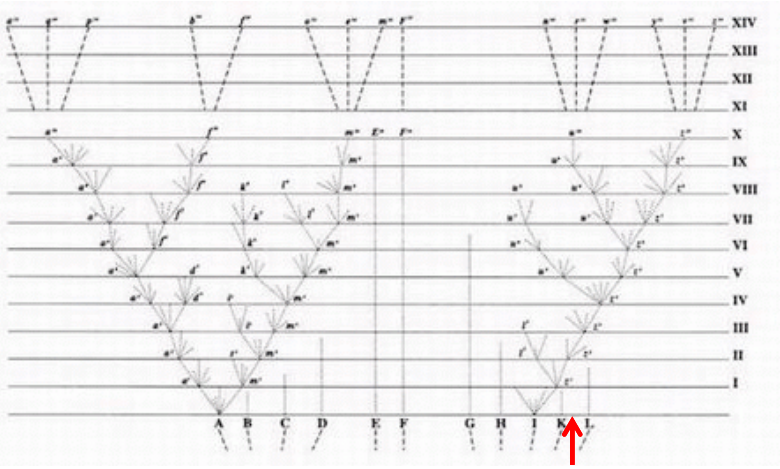
Linnaean taxonomy

1735 (*Species Plantarum*) 1st ed.

http://www.ucmp.berkeley.edu/history/linnaeus.html

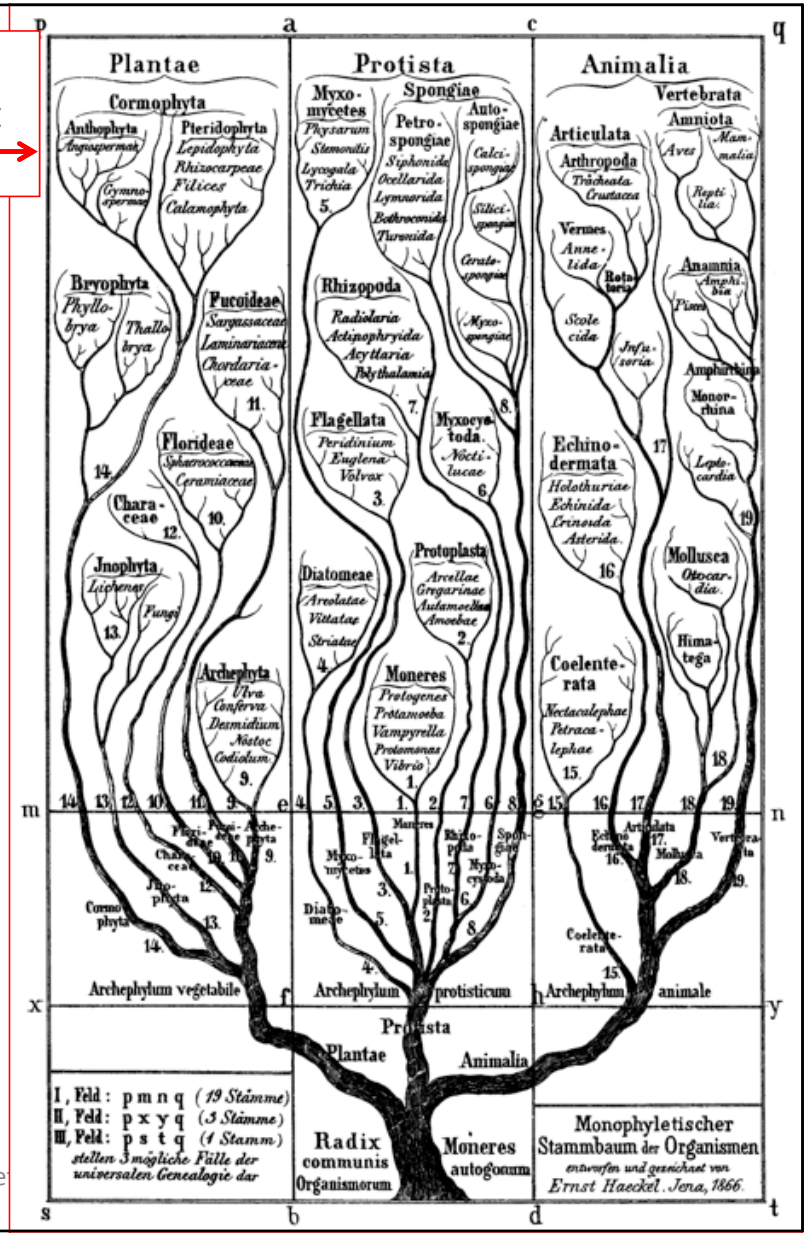


Page from Darwin's notebooks around July 1837 showing his first sketch of an evolutionary tree



Darwin, Charles (1859). On the Origin of Species, pp. 116–117.

Generelle Morphologie der Organismen by Ernst Haeckel (1866)



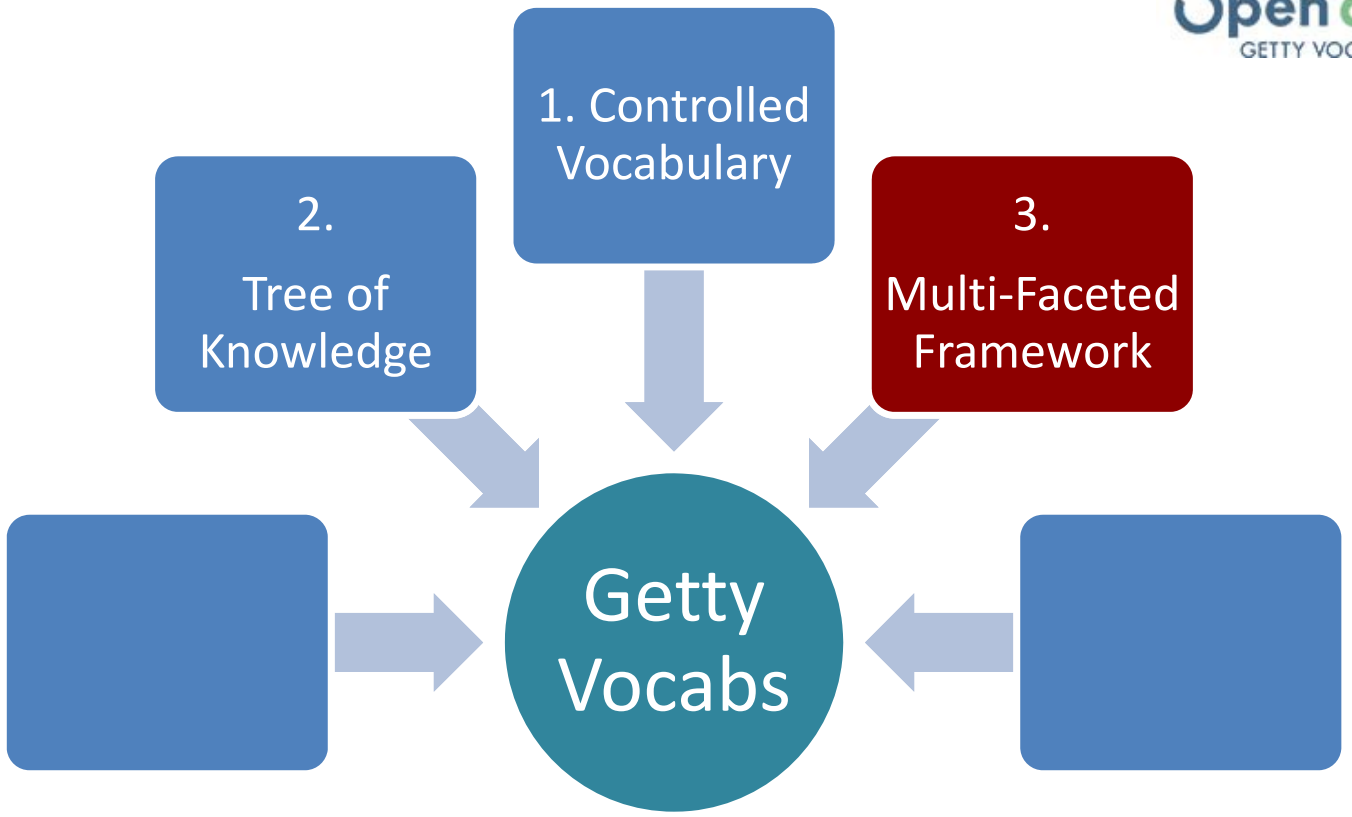
https://en.wikipedia.org/wiki/Tree_of_life_%28biology%29

Marcia Zeng@ Ge

Top of the AAT hierarchies
.... Associated Concepts Facet
..... Associated Concepts (hierarchy name)
.... Physical Attributes Facet
..... Attributes and Properties (hierarchy name)
..... Conditions and Effects (hierarchy name)
..... Design Elements (hierarchy name)
..... Color (hierarchy name)
.... Styles and Periods Facet
..... Styles and Periods (hierarchy name)
.... Agents Facet
..... People (hierarchy name)
..... Organizations (hierarchy name)
..... Living Organisms (hierarchy name)
..... agents (general) [N]
.... Activities Facet
..... Disciplines (hierarchy name)
..... Functions (hierarchy name)
..... Events (hierarchy name)
..... Physical and Mental Activities (hierarc
..... Processes and Techniques (hierarchy
..... activities (general context)
.... Materials Facet
..... Materials (hierarchy name)
.... Objects Facet
..... Built Environment (hierarchy name)
..... Components (hierarchy name)
..... Furnishings and Equipment (hierarchy
..... Object Genres (hierarchy name)
..... Object Groupings and Systems (hierar
..... Visual and Verbal Communication (hie
.... Brand Names Facet
..... Brand Names (hierarchy name)
Top of the AAT hierarchies
.... Agents Facet
..... Living Organisms (hierarchy name)
..... living organisms (entities)
..... Archaea (domain)
..... Bacteria (domain)
..... cyanobacteria
..... Firmicutes (phylum)
..... Eukaryota (domain)
..... Animalia (kingdom)
..... Chromista (kingdom)
..... Fungi (kingdom)
..... Plantae (kingdom)
..... Protista (kingdom)
..... protozoans
..... slime molds
..... <organisms by activity>
..... biofilm
..... bioindicators
..... parasites
..... pests (organisms)
..... <organisms by characteristics>
..... aerobes
..... anaerobes
..... microorganisms
..... <organisms by condition>
..... ephemeral communities (plant or animal communities)
..... fossils
..... Virus (kingdom)
..... wildlife
..... wild animals [N]
Getty Vocabs
=
Tree of Knowledge
11

What is the "Getty Vocabularies"?

(i.e., Why does any dataset need to care about it?)



Ranganathan's Faceted Classification

- developed prior to the existence of computers

PMEST facets:

- **Personality** [P] is best thought of as "the thing itself,"
- **Matter** [M] is the material of which the thing is composed,
- **Energy** [E] is the action performed on or by the thing,
- **Space** [S] is where the action takes place,
- **Time** [T] is when it takes place.

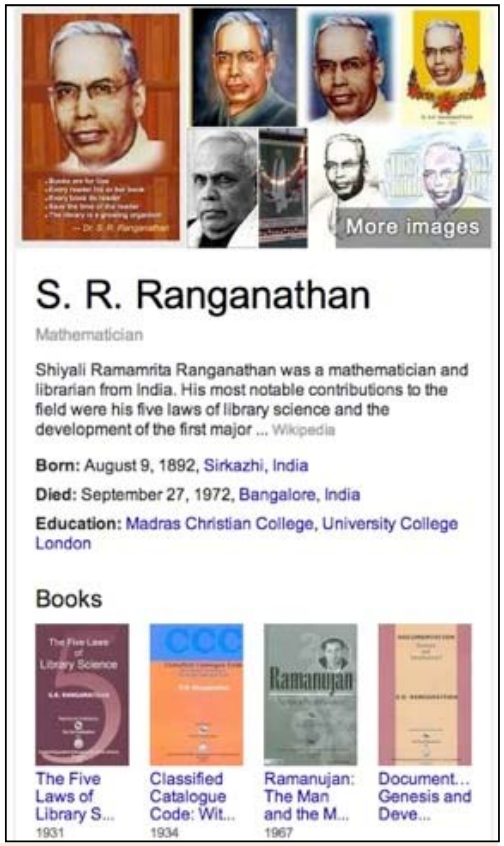
WHO

WHAT

HOW

WHERE

WHEN



Colon Classification 1933-

Synthesis power

'What distinguishes the universe of current knowledge is that it is a dynamical continuum. It is ever growing; new branches may stem from any of its infinity of points at any time; they are unknowable at present. They cannot therefore be enumerated here and now; nor can they be anticipated, their filiations can be determined only after they appear'' (Ranganathan, 1951).

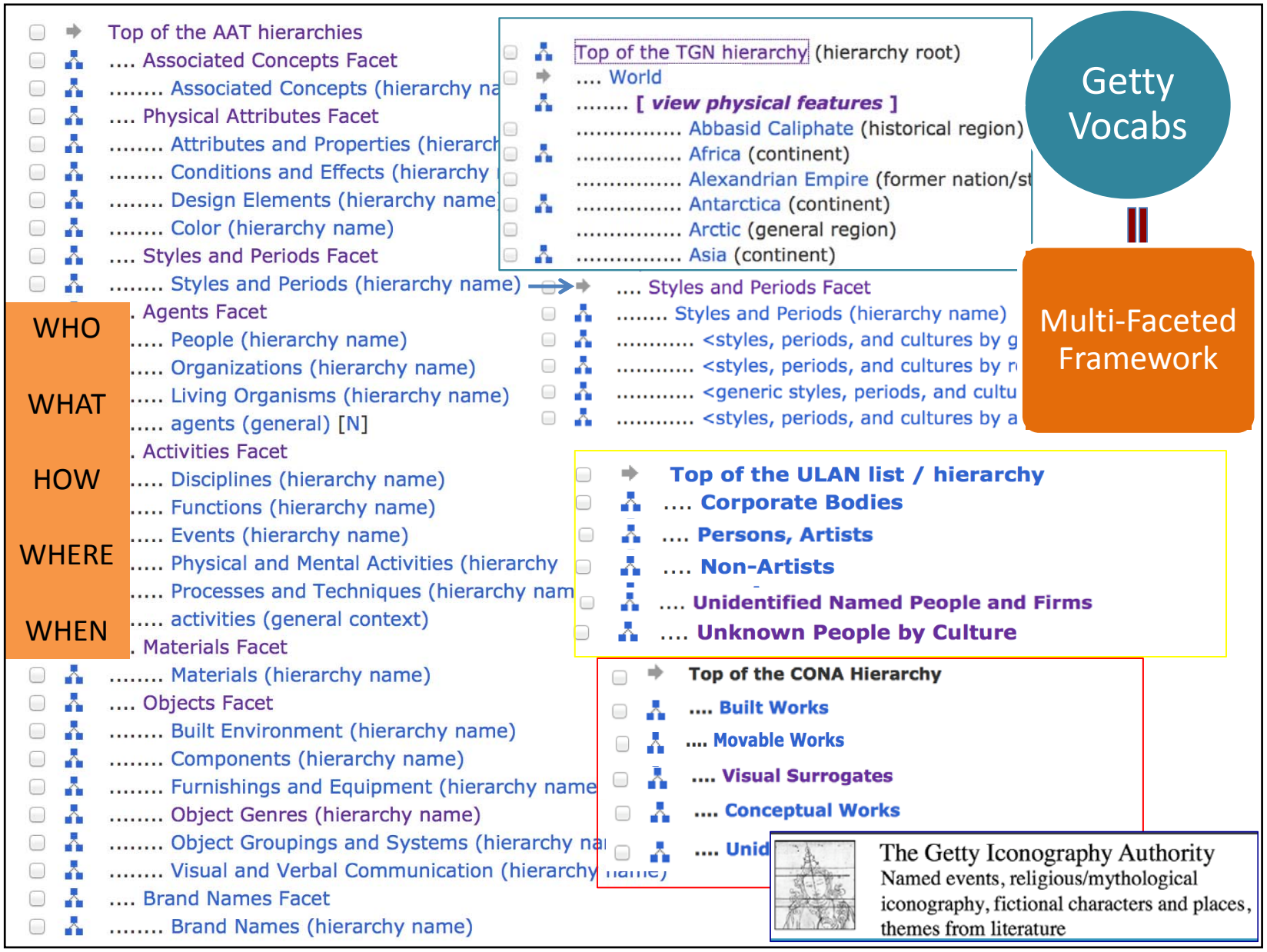
EXPLAINNING THE FACETED APPROACH

Applications of Faceted Structures

Many types of information tools and systems have been designed from faceted principles.

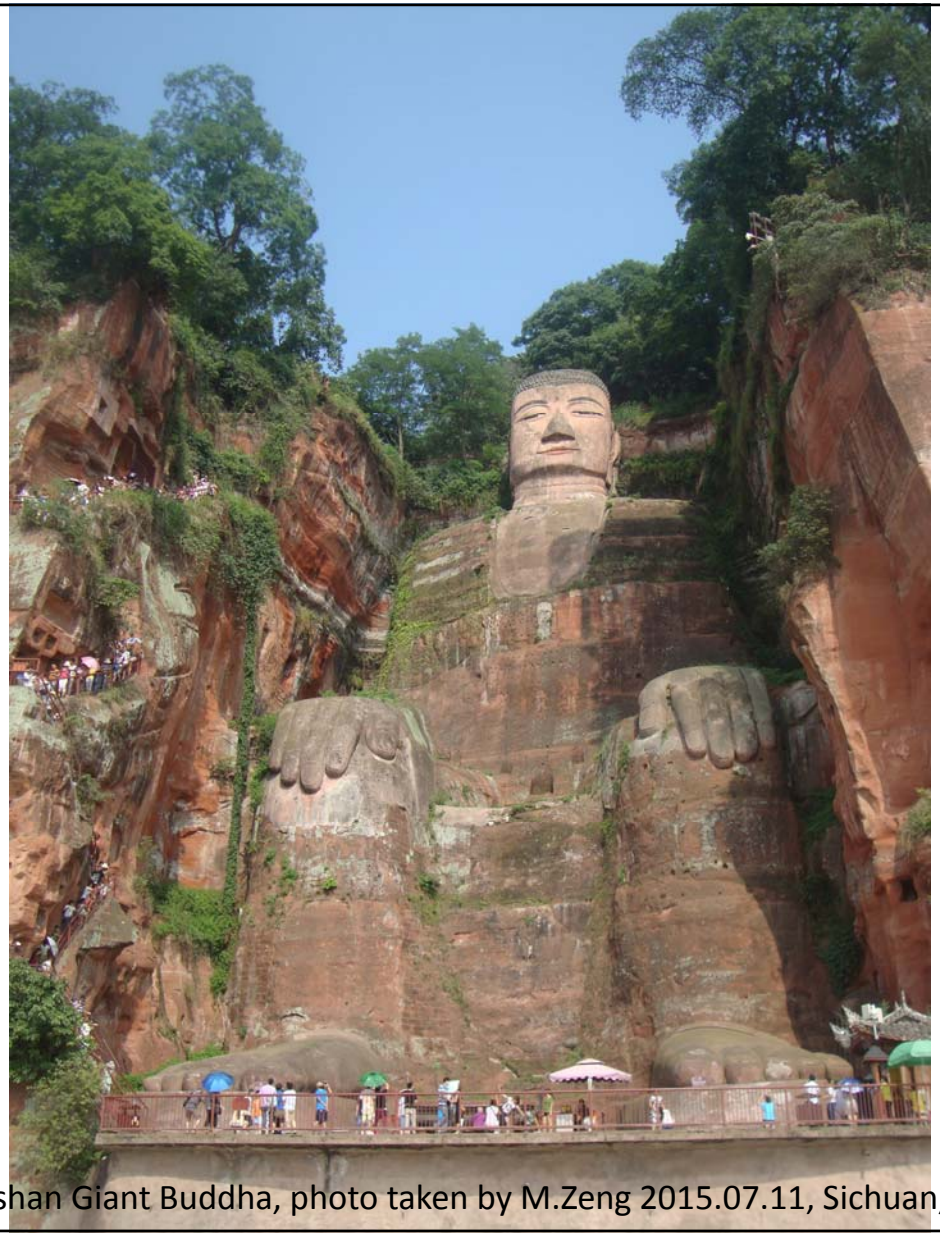
- Classification schemes
  - *Universal Decimal Classification (UDC)*
  - *Colon Classification*
- Faceted thesauri
  - *Art and Architecture Thesaurus (AAT)*
  - *Thesaurofacets*
  - Library of Congress' new vocabularies
- Computerized indexing systems
  - E.g., PRECIS, POPSI
- Expert systems
- Information architecture
  - websites
  - data visualization
- Ontologies

15


Top of the AAT hierarchies
.... Associated Concepts Facet
..... Associated Concepts (hierarchy na
.... Physical Attributes Facet
..... Attributes and Properties (hierarch
..... Conditions and Effects (hierarchy
..... Design Elements (hierarchy name
..... Color (hierarchy name)
.... Styles and Periods Facet
..... Styles and Periods (hierarchy name)
WHO
. Agents Facet
..... People (hierarchy name)
..... Organizations (hierarchy name)
WHAT
..... Living Organisms (hierarchy name)
..... agents (general) [N]
. Activities Facet
HOW
..... Disciplines (hierarchy name)
..... Functions (hierarchy name)
..... Events (hierarchy name)
WHERE
..... Physical and Mental Activities (hierarchy
..... Processes and Techniques (hierarchy nam
WHEN
..... activities (general context)
. Materials Facet
..... Materials (hierarchy name)
.... Objects Facet
..... Built Environment (hierarchy name)
..... Components (hierarchy name)
..... Furnishings and Equipment (hierarchy name
..... Object Genres (hierarchy name)
..... Object Groupings and Systems (hierarchy na
..... Visual and Verbal Communication (hierarchy
.... Brand Names Facet
..... Brand Names (hierarchy name)
Top of the TGN hierarchy (hierarchy root)
.... World
..... [view physical features]
..... Abbasid Caliphate (historical region)
..... Africa (continent)
..... Alexandrian Empire (former nation/st
..... Antarctica (continent)
..... Arctic (general region)
..... Asia (continent)
.... Styles and Periods Facet
..... Styles and Periods (hierarchy name)
..... <styles, periods, and cultures by g
..... <styles, periods, and cultures by r
..... <generic styles, periods, and cultu
..... <styles, periods, and cultures by a
Top of the ULAN list / hierarchy
.... Corporate Bodies
.... Persons, Artists
.... Non-Artists
.... Unidentified Named People and Firms
.... Unknown People by Culture
Top of the CONA Hierarchy
.... Built Works
.... Movable Works
.... Visual Surrogates
.... Conceptual Works
.... Unid
The Getty Iconography Authority
Named events, religious/mythological iconography, fictional characters and places, themes from literature
Getty Vocabs
Multi-Faceted Framework

Leshan Giant Buddha Scenic Area

- a UNESCO World Heritage Site

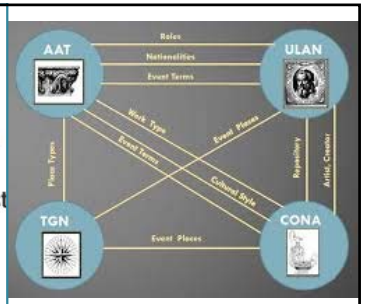


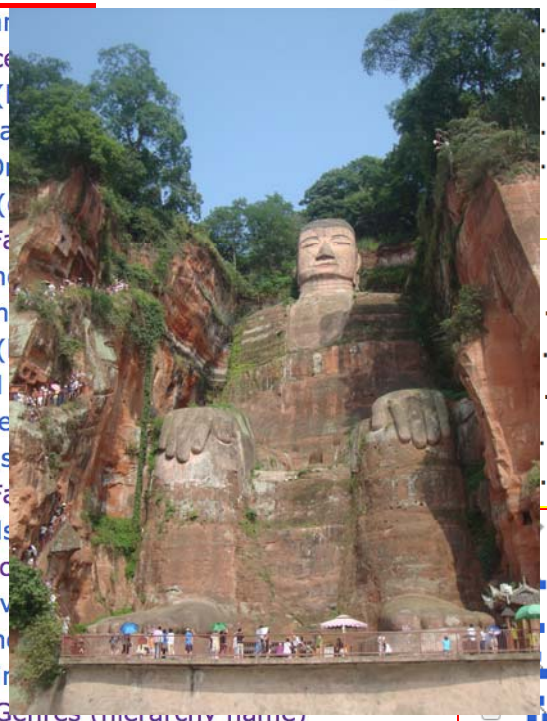
- 71-metre (233 ft) tall stone statue,
- built during the Tang Dynasty (618–907),
- depicting Maitreya (彌勒菩薩), a bodhisattva, (a future Buddha).

Leshan Giant Buddha, photo taken by M.Zeng 2015.07.11, Sichuan, China

Top of the AAT hierarchies
.... Associated Concepts Facet
..... Associated Concepts (hierarchy na
.... Physical Attributes Facet
..... Attributes and Properties (hierarch
..... Conditions and Effects (hierarchy
..... Design Elements (hierarchy name
..... Color (hierarchy name)
.... Styles and Periods Facet
..... Styles ar
.... Agents Face
..... People (
..... Organiza
..... Living Or
..... agents (
.... Activities Fa
..... Disciplin
..... Function
..... Events (
..... Physical
..... Processe
..... activities
.... Materials Fa
..... Materials
.... Objects Fac
..... Built Env
..... Compon
..... Furnishi
..... Object G
..... Object Groupings and Systems (hierarchy na
..... Visual and Verbal Communication (hierarchy
.... Brand Names Facet
..... Brand Names (hierarchy name)

Top of the TGN hierarchy (hierarchy root)
.... World
..... [*view physical features*]
..... Abbasid Caliphate (historical region)
..... Africa (continent)
..... Alexandrian Empire (former nation/st
..... Antarctica (continent)
..... Arctic (general region)
..... Asia (continent)



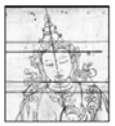


How cultural objects (and their images) can be researched /studied/ exhibited/displayed/ linked/ searched/ browsed/shared/ liked/…?

--Getty Vocabs together provides a multi-faceted framework for organizing data and information for them.

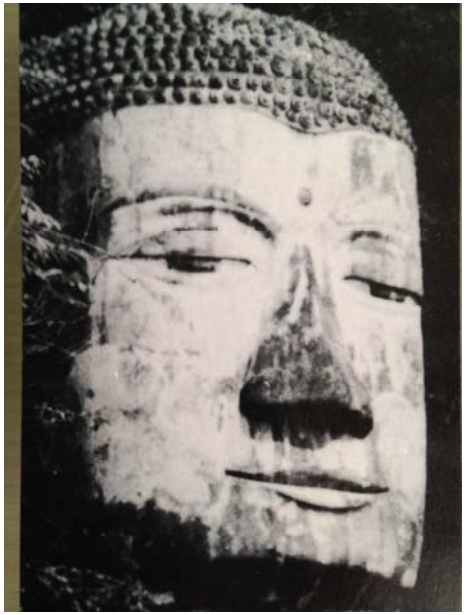
.... **Non-Artists**
.... **Unidentified Named People and Firms**
.... **Unknown People by Culture**

Top of the CONA Hierarchy
.... **Built Works**
.... **Movable Works**
.... **Visual Surrogates**
.... **Conceptual Works**
.... **Unid**

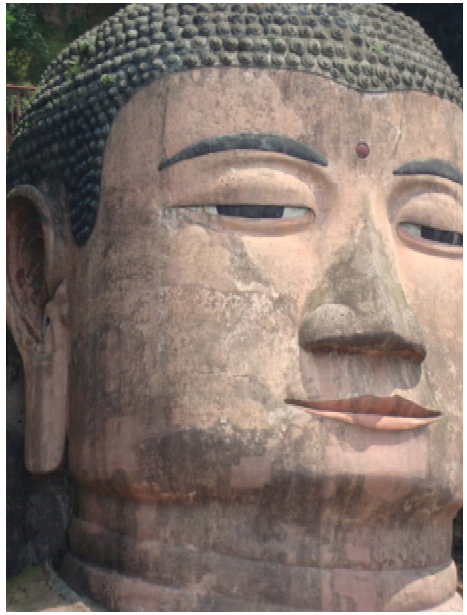


The Getty Iconography Authority
Named events, religious/mythological iconography, fictional characters and places, themes from literature

1962



1963



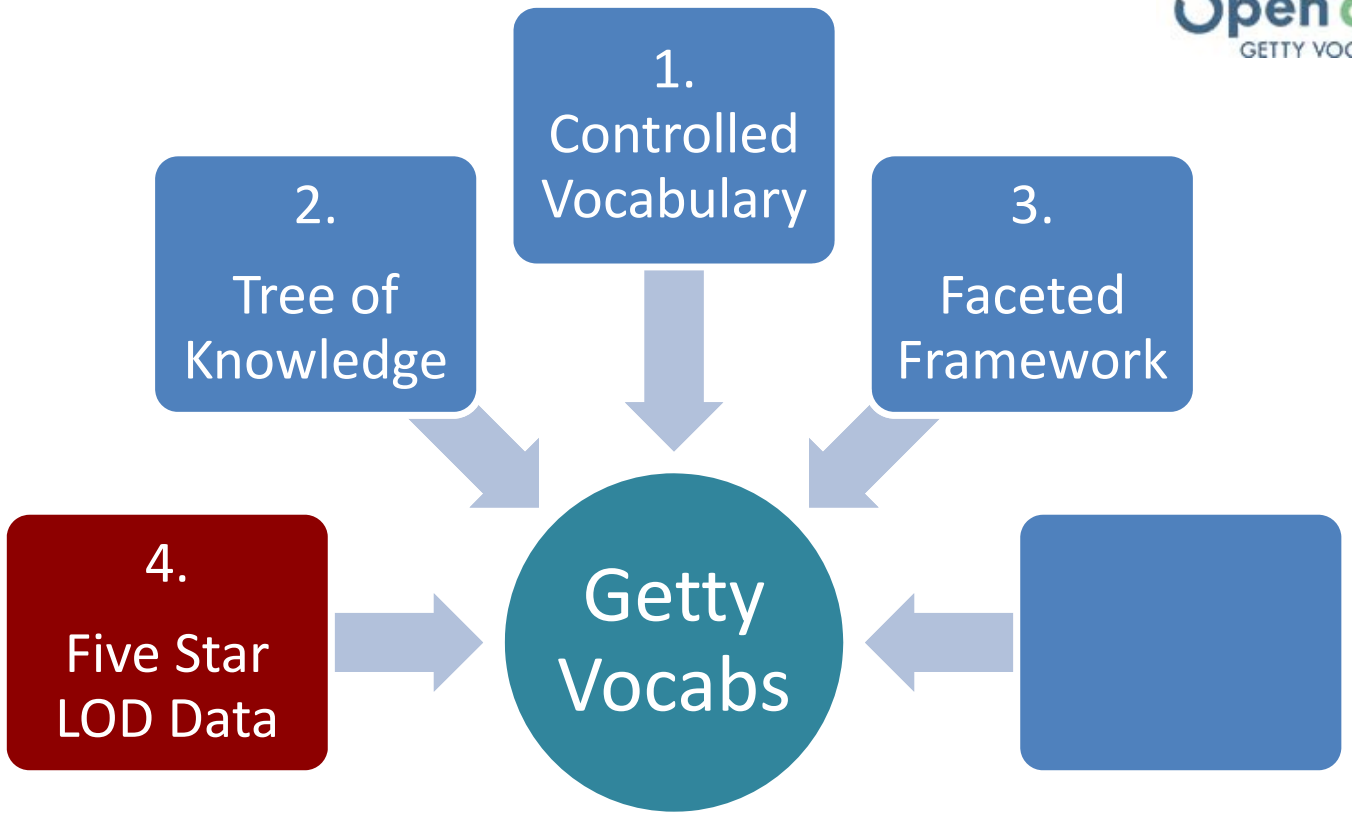
2015

1959-1961: Three Years
of Natural Disasters

Images from a set of postcards.

What is the "Getty Vocabularies"?

(i.e., Why does any dataset need to care about it?)



Art & Architecture Thesaurus (AAT)'s Path to LOD

- ☑ 1970s started
- ☑ 1983 @ the Getty
- ☑ 1990, 1994 Published (hardcopy and e-version)

- ☑ 2011.07 SKOSifying pilot study

- ☑ 2013 ontology
- ☑ 2014.02 published as LOD

Controlled vocabulary

↓

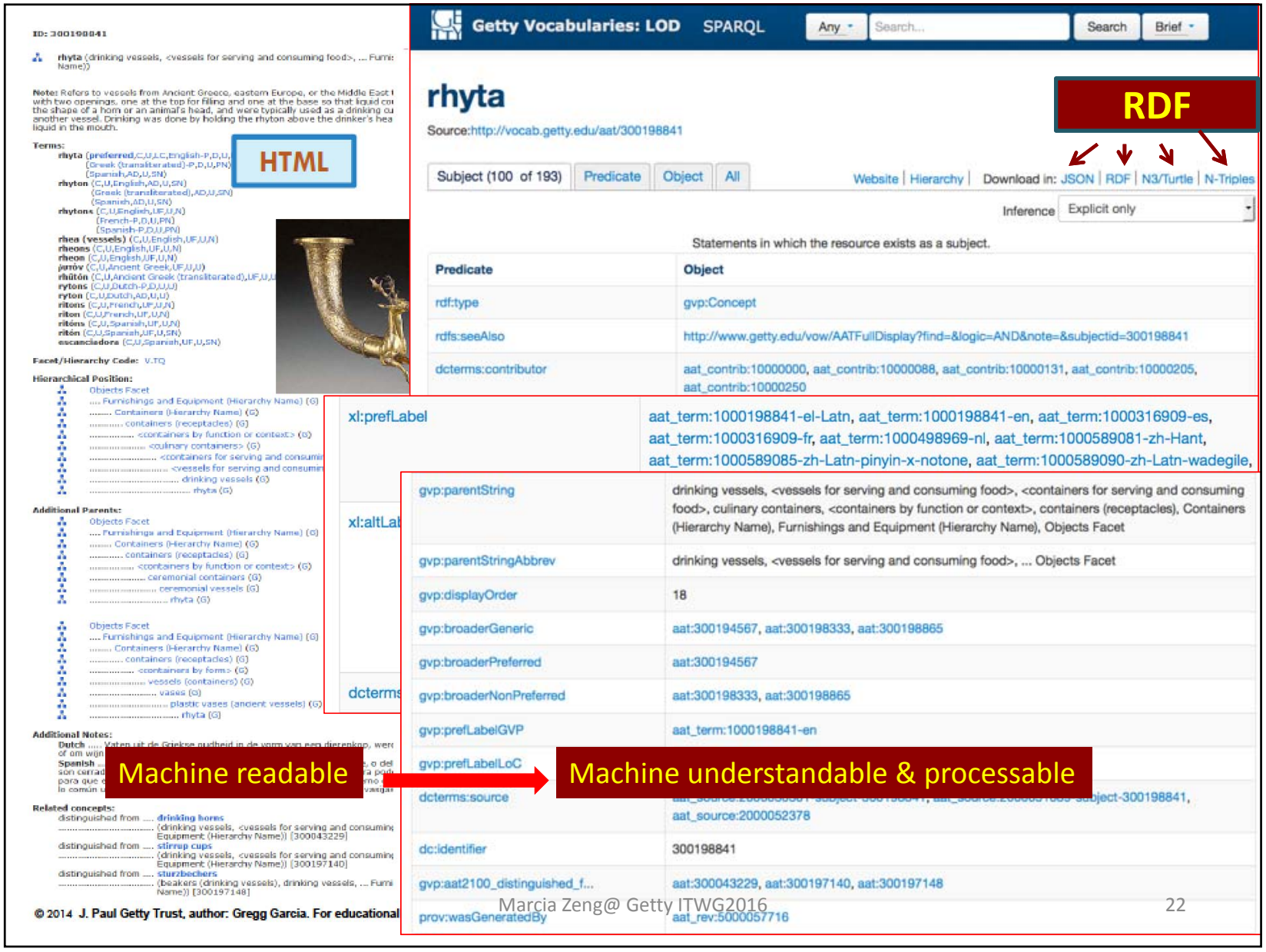
SKOSified value vocabulary

↓

LOD dataset, a knowledge base

AAT 2016.08-01： concepts： 45077； terms： 357409*

*Results based on the query links at https://en.wikipedia.org/wiki/Art_%26_Architecture_Thesaurus for counting 'concepts' and 'terms'.

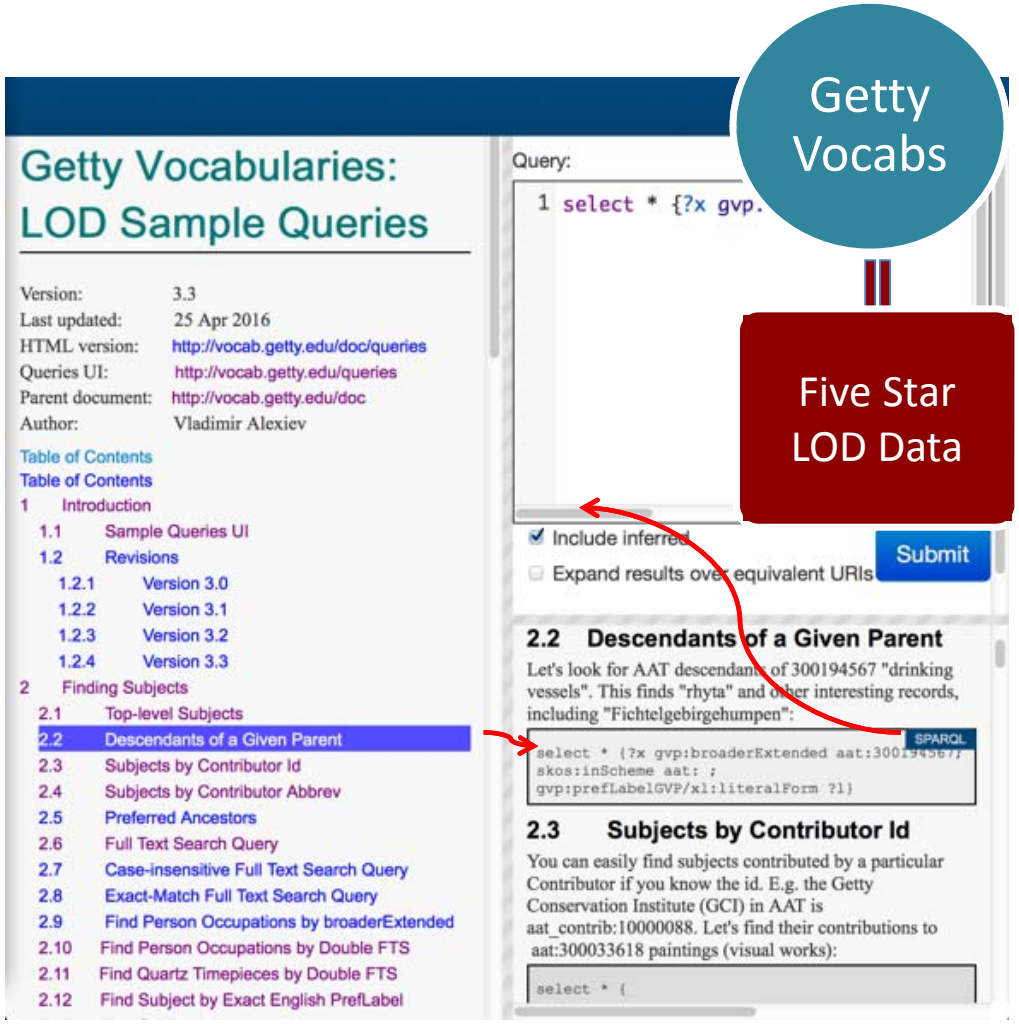

HTML
Getty Vocabularies: LOD SPARQL
Any
Search...
Search
Brief
rhyta
Source:http://vocab.getty.edu/aat/300198841
RDF
Subject (100 of 193)
Predicate
Object
All
Website | Hierarchy |
Download in: JSON | RDF | N3/Turtle | N-Triples
Inference
Explicit only
Statements in which the resource exists as a subject.
Predicate
Object
rdf:type
gvp:Concept
rdfs:seeAlso
http://www.getty.edu/vow/AATFullDisplay?find=&logic=AND&note=&subjectid=300198841
dcterms:contributor
aat_contrib:10000000, aat_contrib:10000088, aat_contrib:10000131, aat_contrib:10000205, aat_contrib:10000250
xl:prefLabel
aat_term:1000198841-el-Latn, aat_term:1000198841-en, aat_term:1000316909-es, aat_term:1000316909-fr, aat_term:1000498969-nl, aat_term:1000589081-zh-Hant, aat_term:1000589085-zh-Latn-pinyin-x-notone, aat_term:1000589090-zh-Latn-wadegile,
gvp:parentString
drinking vessels, <vessels for serving and consuming food>, <containers for serving and consuming food>, culinary containers, <containers by function or context>, containers (receptacles), Containers (Hierarchy Name), Furnishings and Equipment (Hierarchy Name), Objects Facet
gvp:parentStringAbbrev
drinking vessels, <vessels for serving and consuming food>, ... Objects Facet
gvp:displayOrder
18
gvp:broaderGeneric
aat:300194567, aat:300198333, aat:300198865
gvp:broaderPreferred
aat:300194567
gvp:broaderNonPreferred
aat:300198333, aat:300198865
gvp:prefLabelGVP
aat_term:1000198841-en
gvp:prefLabelLoC
dcterms:source
aat_source:2000052378
dc:identifier
300198841
gvp:aat2100_distinguished_f...
aat:300043229, aat:300197140, aat:300197148
prov:wasGeneratedBy
aat_rev:5000057716
Machine readable
Machine understandable & processable

- AAT release: 2014.02
- TGN release: 2014.08
- ULAN released: 2015.04
- CONA: [2016.01]

ODC BY 1.0

- Ontology version 3.3

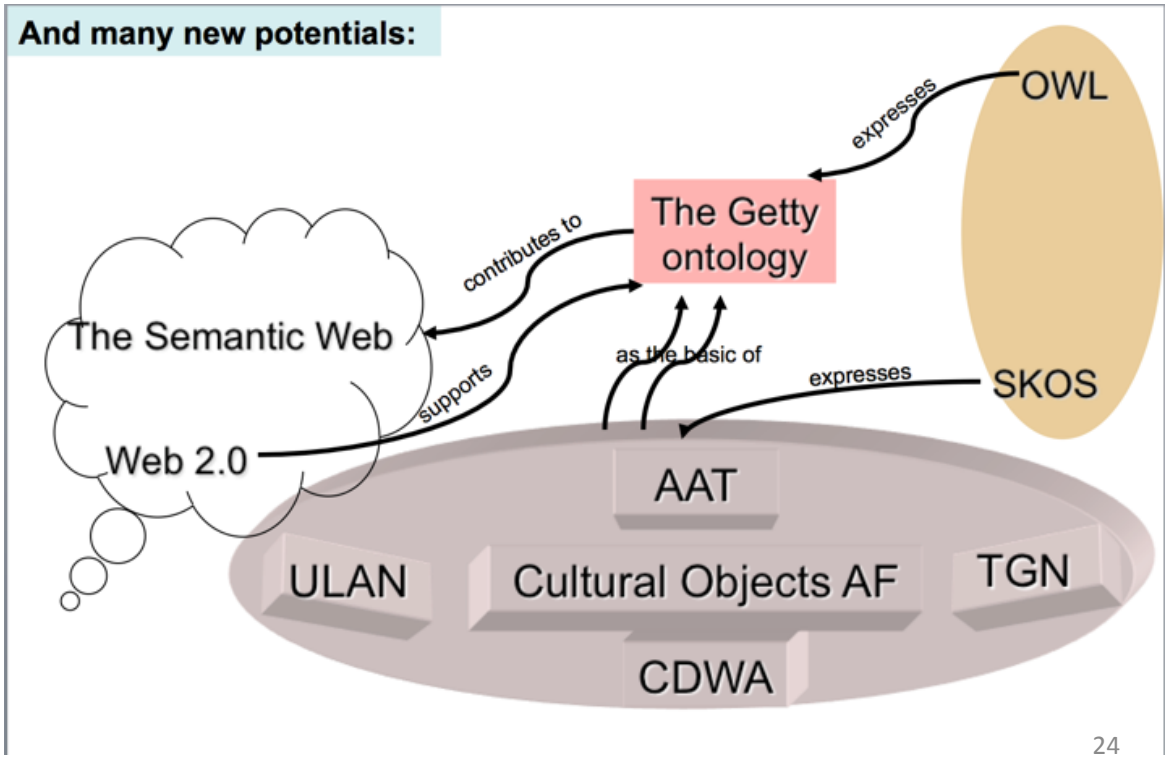
In addition to SKOS & SKOS-XL，it uses properties from other RDF vocabularies：FOAF, PROV, Schema, DC, DCT, ISO, RDF, RDFs, OWL, BIBO, WGS, XSD…



More at https://share.getty.edu/display/ITSLODV/AAT+Semantic+Representation

http://vocab.getty.edu/queries#Finding_Subjects

Looks like the imagination has become a reality!



- Zeng, M.L. 2008-03-11. Discussions: The Semantic Web

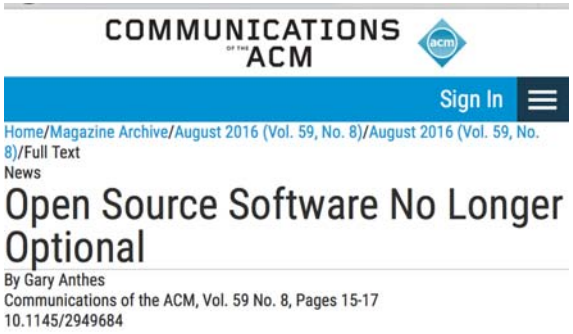
Note: "Open" is not simple

Using Open Source Software (OSS) as our example:

Anthes, Gary. 2016. "Open Source Software No longer Optional" *Communications of the ACM*. Aug. 2016, 59(8): 15- 17.

Open development and sharing of software gained widespread acceptance 15 years ago, and the practice is accelerating.

> "[Keepers, GitHub's head of open source software:] 'We are seeing companies treating open source launches like product launches. They want to make a big splash, but they want to make sure there is support for the project after the launch.'" (Anthes, 2016, p.17)


COMMUNICATIONS OF THE ACM

Sign In

Home/Magazine Archive/August 2016 (Vol. 59, No. 8)/August 2016 (Vol. 59, No. 8)/Full Text

News

Open Source Software No Longer Optional

By Gary Anthes

Communications of the ACM, Vol. 59 No. 8, Pages 15-17

10.1145/2949684

-- *Communications of the ACM*. Aug. 2016, 59(8): 15- 17.

http://m.cacm.acm.org/magazines/2016/8/205050-open-source-software-no-longer-optional/fulltext

Using Open Source Software (OSS) as our ex

Linux

a Unix-like computer operating system (OS) assembled under the model of free and open-source software development and distribution

- 1991, s
  compu
- Today,
- Tens o
  server
- e.g., A
  superc

" 'We are seeing companies treating open source launches like product launches. They want to make a big splash, but they want to make sure there is support for the project after the launch.'" (Anthes, 2016, p.17)

"Open" requires sustained efforts and strong supports.

OpenSSL

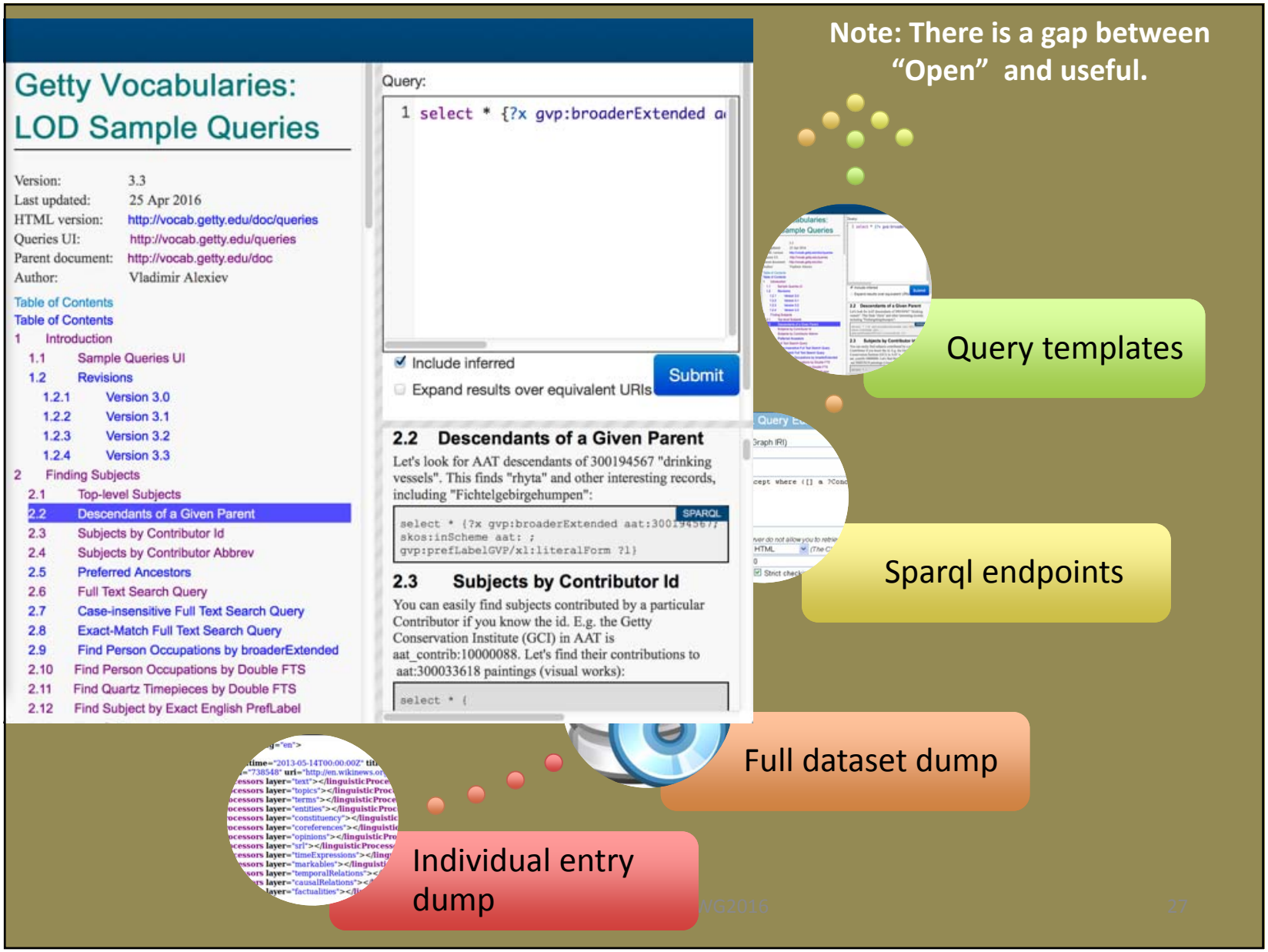
a software library to be used in applications that need to secure communications against eavesdropping or need to ascertain the identity of the party at the other end.

- As of 2014 two thirds of all webservers use OpenSSL
- Wasn't a well-funded consortium, (the project has a budget of less than $1 million a year and relies in part on donations.)
- The management team consists of four Europeans. The entire development group consists of 11 members, out of which 10 are volunteers; there is only one full-time employee,
- In 2014 the bug left an estimated 500,000 computers vulnerable to breaches of cryptographic security.

GitHub

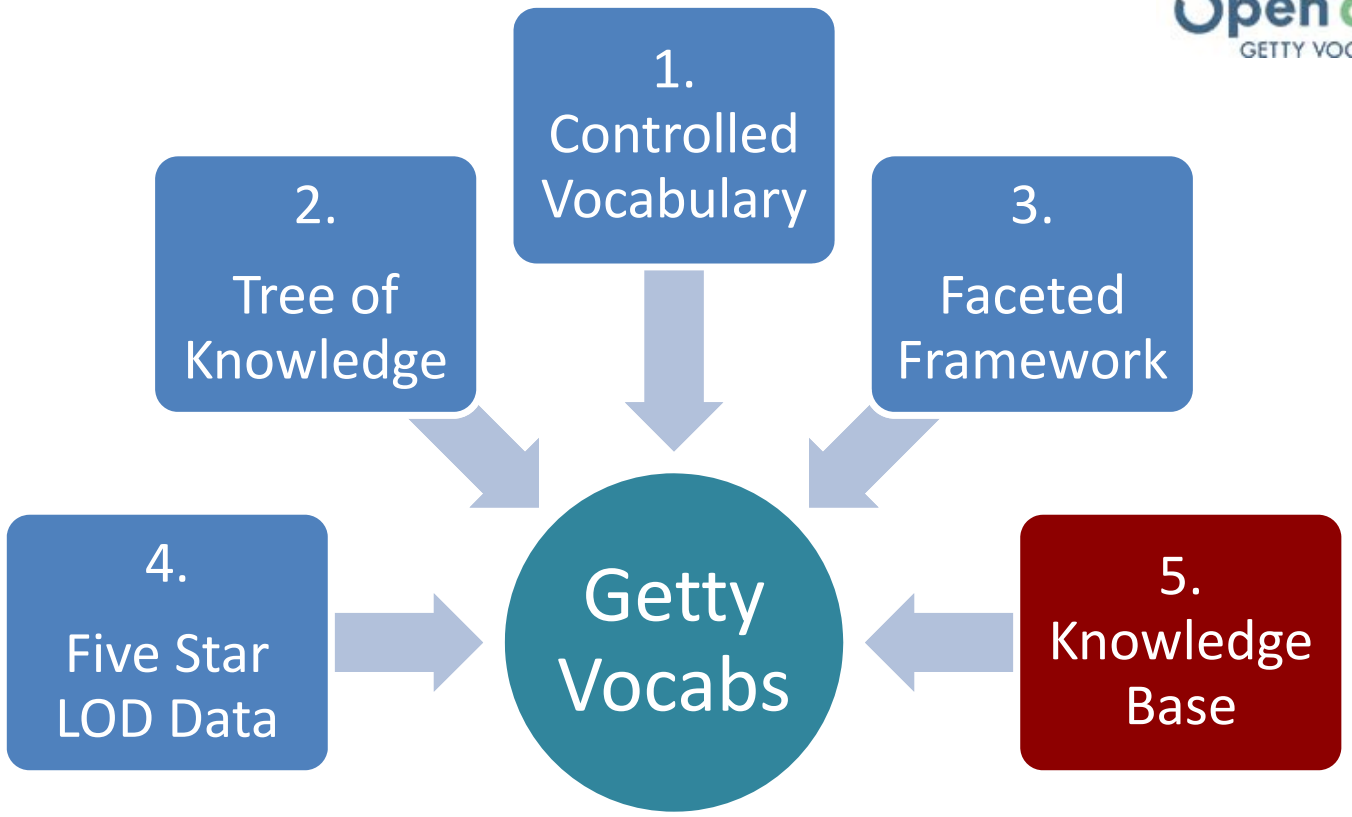
a web-based Git (software) repository hosting service

- the company GitHub has become the go-to place for developers and users of open software
- Users: large companies such as Apple, Google, Microsoft
- Users: thousands of start-ups
- Host 31 million open sirce projects used by 12 million developers.
- As of April 2016, GitHub reports having more than 14 million users and more than 35 million repositories, making it the largest host of source code in the world.


Getty Vocabularies:
LOD Sample Queries
Version: 3.3
Last updated: 25 Apr 2016
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
1.2.3 Version 3.2
1.2.4 Version 3.3
2 Finding 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
2.7 Case-insensitive Full Text Search Query
2.8 Exact-Match Full Text Search Query
2.9 Find Person Occupations by broaderExtended
2.10 Find Person Occupations by Double FTS
2.11 Find Quartz Timepieces by Double FTS
2.12 Find Subject by Exact English PrefLabel
Query:
1 select * {?x gvp:broaderExtended a
Include inferred
Expand results over equivalent URIs
Submit
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":
SPARQL
select * {?x gvp:broaderExtended aat:300194567;
skos:inScheme aat: ;
gvp:prefLabelGVP/xl:literalForm ?l}
2.3 Subjects by Contributor Id
You can easily find subjects contributed by a particular Contributor if you know the id. E.g. the Getty Conservation Institute (GCI) in AAT is aat_contrib:10000088. Let's find their contributions to aat:300033618 paintings (visual works):
select * {
Note: There is a gap between "Open" and useful.
Query templates
Sparql endpoints
Full dataset dump
Individual entry dump

What is the "Getty Vocabularies"?

(i.e., Why does any dataset need to care about it?)



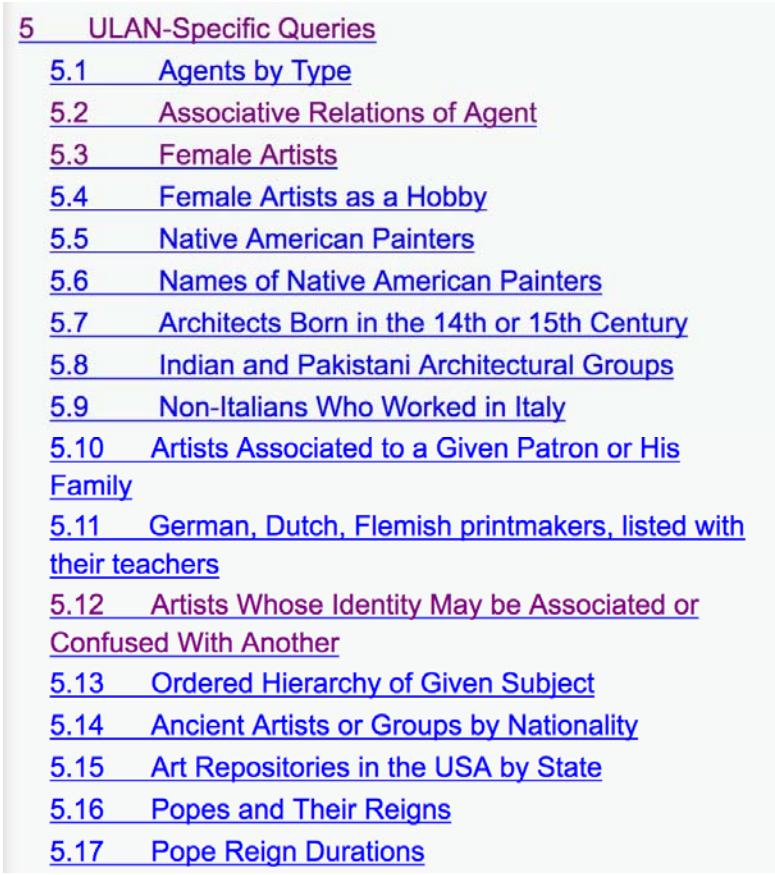
As knowledge bases of research

LOD KOS can be used for

- obtaining special graphs or datasets for very complicated questions, and
- revealing unknown relationships

  e.g.,

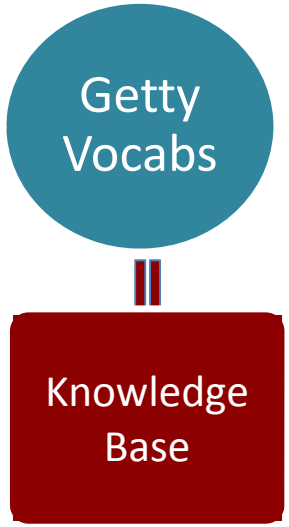
  - associative relations of agent (people or organization),
  - places by type within a geo-bounding box,
  - scientific names not in English or Latin,
  - …



Getty Vocabularies: LOD Sample Queries

4 TGN-Specific Queries
- 4.1 Places by Type
- 4.2 Places, with English or GVP Label
- 4.3 Places by Direct and Hierarchical Type
- 4.4 Breakdown of Sovereign States by Type
- 4.5 Inhabited Places That Were Sovereign States
- 4.6 Places by Type and Parent Place
- 4.7 Places by Type, with placeTypePreferred
- 4.8 Places by Triple FTS
- 4.9 Places by FTS Parents
- 4.10 Capitals by Association
- 4.11 Members of the European Union
- 4.12 Members of the United Nations
- 4.13 Geo Chart with sgvizler
- 4.14 Column Chart with sgvizler
- 4.15 Countries and Capitals By Type and Containment
- 4.16 Places by Coordinate Bounding Box
- 4.17 Places Within Bounding Box
- 4.18 Places by Type Within Bounding Box
- 4.19 Places Outside Bounding Box (Overseas Possessions)
- 4.20 Places Nearby Each Other

http://vocab.getty.edu/queries#Top-level_Subjects

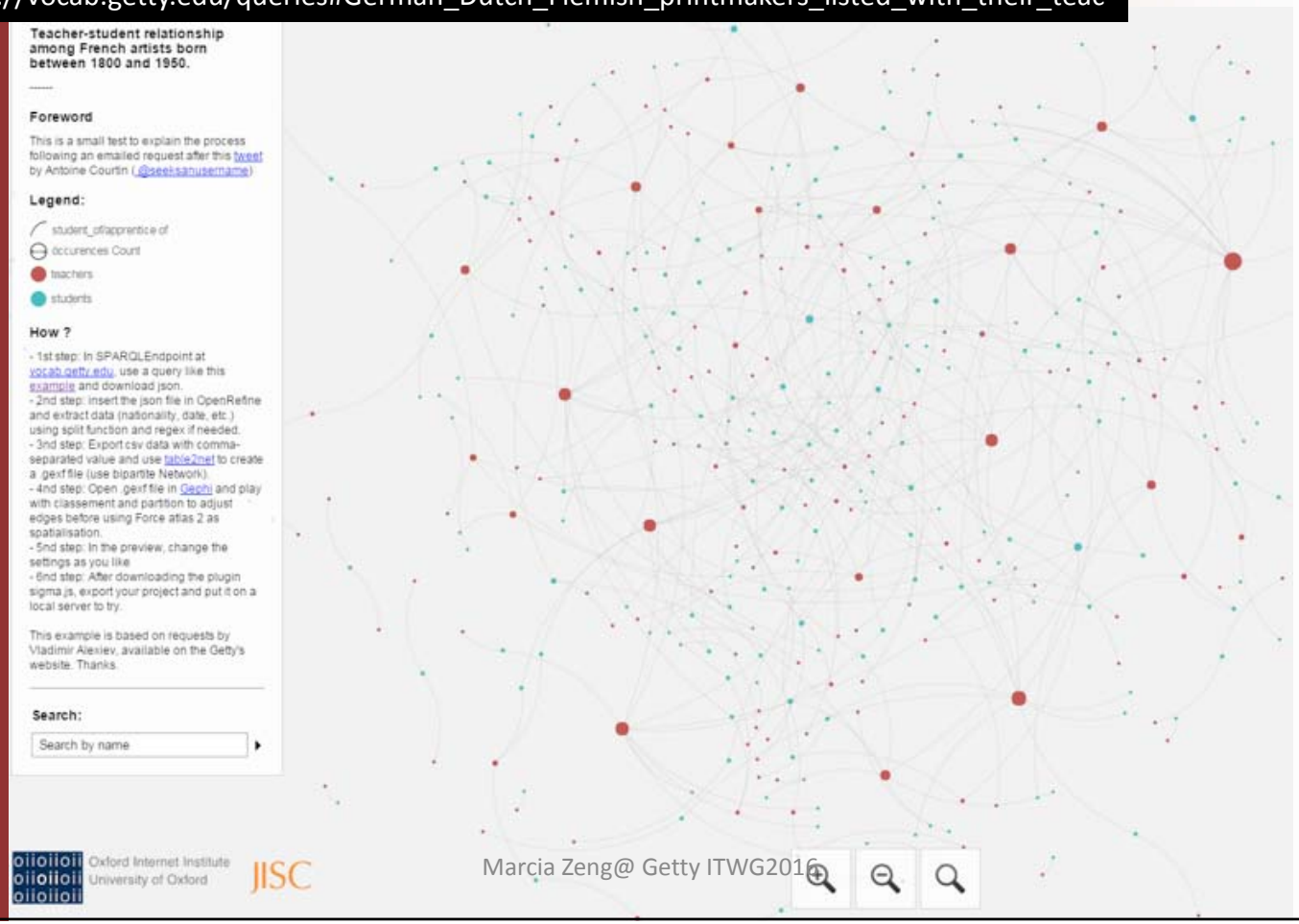


- obtaining special graphs or datasets for very complicated questions, and
- revealing unknown relationships

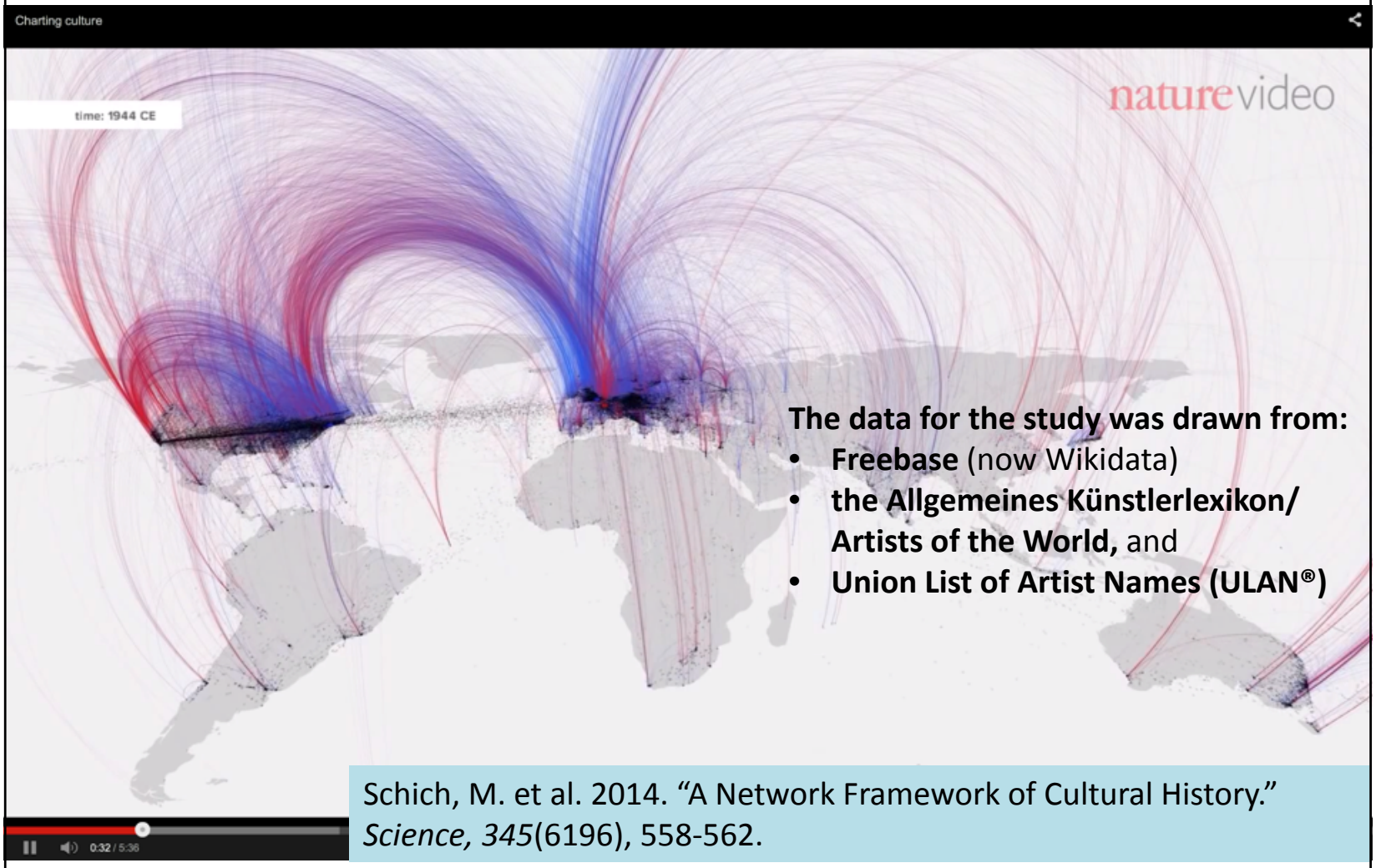
Example: Getty LOD Vocab be the foundation of a network analysis

Teacher-student relationship among French artists born between 1800 and 1950.

query http://vocab.getty.edu/queries#German_Dutch_Flemish_printmakers_listed_with_their_teachers



Nature Video. (2014, July 31). **Charting culture.** https://www.youtube.com/watch?v=4gIhRkCcD4U



The data for the study was drawn from:

- **Freebase** (now Wikidata)
- **the Allgemeines Künstlerlexikon/ Artists of the World,** and
- **Union List of Artist Names (ULAN®)**

Schich, M. et al. 2014. "A Network Framework of Cultural History." *Science, 345*(6196), 558-562.

When the “Getty Vocabularies”
is a 5-star Data,
it enables others to become 5-star too

I. For Vocab Creators/Managers

Getty Vocabs

1. As the resources of
   - creating, maintaining, enriching, extending, and
   - translating a controlled vocabulary

- 2. As the vocabulary management facility

II. For Data Producers & Providers

Getty Vocabs

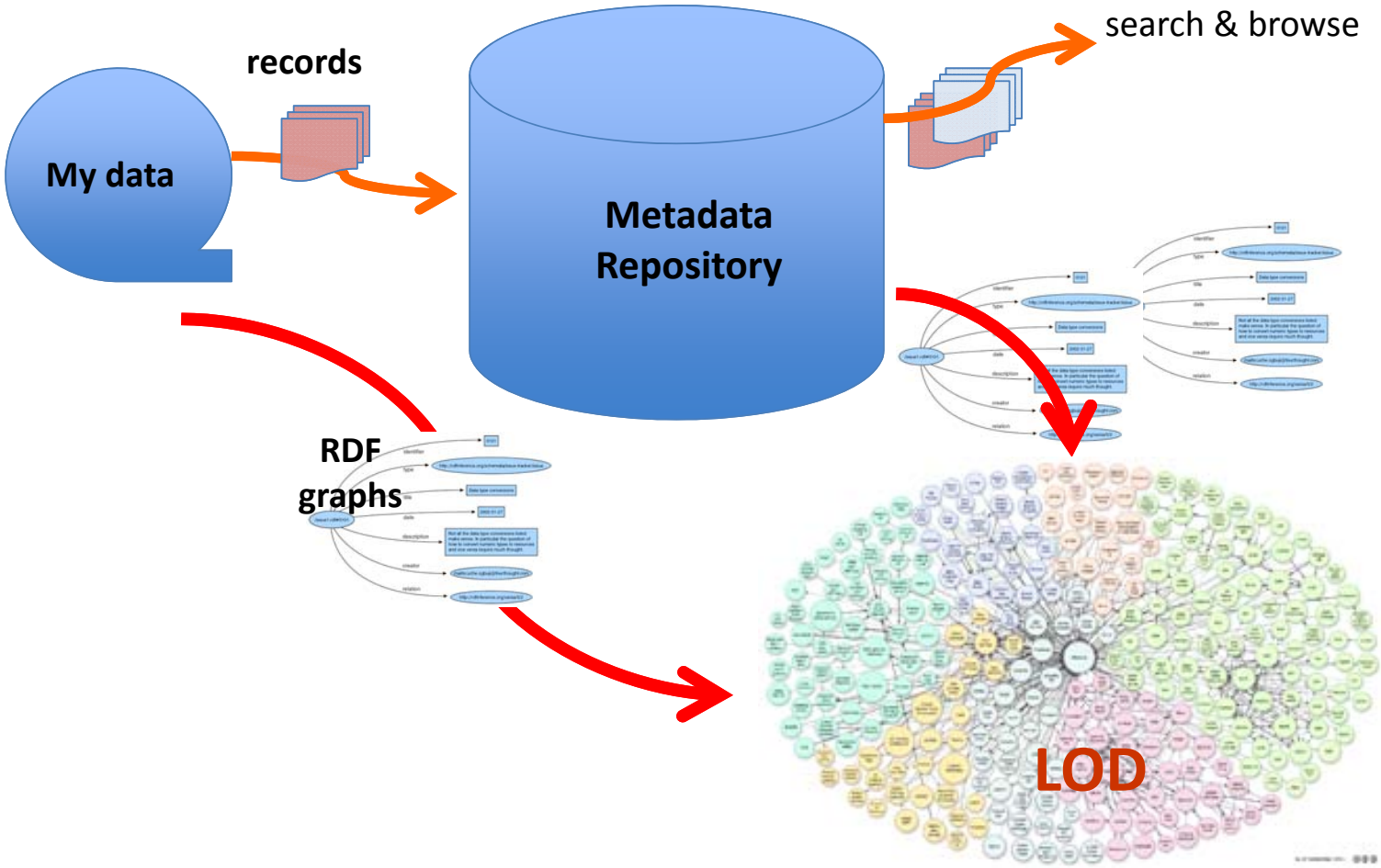
Transforming databases to LOD Datasets

1. Enable owners of structured data to convert and publish their metadata under the LOD principles
   i.e., use HTTP URIs/IRIs as names of things

2. Enhance semantic consistency and interoperability

3. Increase the findability of their data.

Output your data



& Connecting your data to other LOD datasets

Use LOD KOS APIs
--mapping outsiders

nomisma.org/id/ephesus

ephesus (nmo:Mint)

skos:prefLabel	Efese *(af)*, ኤፌሶን *(am)*, أفسس *(ar)*, Efes *(az)*, Горад Эфес *(be)*, Ефес *(bg)*, Efesos *(br)*, Efes *(bs)*, Efes *(ca)*, Efez *(cs)*, Effesus *(cy)*, Efesos *(da)*, Ephesos *(de)*, Έφεσος *(el)*, Ephesus *(en)*, Efeso *(eo)*, Éfeso *(es)*, Efeso *(eu)*, افسس *(fa)*, Efesos *(fi)*, Efesus *(fo)*, Ephèse *(fr)*, Éfeso *(gl)*, אפסוס *(he)*, Efez *(hr)*, Epheszosz *(hu)*, Եփեսոս *(hy)*, Ephesos *(id)*, Efesos *(is)*, Efeso *(it)*, エフェソス *(ja)*, Efesus *(jv)*, ეფესო *(ka)*, 에페소스 *(ko)*, Ephesus *(la)*, Efesas *(lt)*, Efesa *(lv)*, Ефес *(mk)*, इफेसूस *(mr)*, Efesus *(ms)*, Efesos *(nb)*, Efese *(nl)*, Efesos *(nn)*, Efez *(pl)*, Éfeso *(pt)*, Efesos *(ro)*, Эфес *(ru)*, Efes *(sh)*, Efez *(sk)*, Efez *(sl)*, Ефес *(sr)*, Efesos *(sv)*, Efeso *(sw)*, எபேசஸ் *(ta)*, เอฟิซัส *(th)*, Efeso *(tl)*, Efes *(tr)*, Ефес *(uk)*, افسس *(ur)*, Efes *(uz)*, Ephesus *(vi)*, 以弗所 *(zh)*
skos:definition	The mint at the ancient site of Ephesus in Ionia. *(en)*
dcterms:isPartOf	http://nomisma.org/id/greek_numismatics
geo:location	http://nomisma.org/id/ephesus#this
rdf:type	skos:Concept
skos:altLabel	Ephesos *(en)*
skos:broader	http://nomisma.org/id/ionia
skos:closeMatch	http://collection.britishmuseum.org/id/place/x107267
skos:closeMatch	http://dbpedia.org/resource/Ephesus
skos:closeMatch	http://gazetteer.dainst.org/app/#!/show/2287728
skos:closeMatch	http://pleiades.stoa.org/places/599612
skos:closeMatch	http://viaf.org/viaf/156158309
skos:closeMatch	http://vocab.getty.edu/tgn/7002499
skos:closeMatch	http://www.geonames.org/7522155
skos:closeMatch	http://www.wikidata.org/entity/Q47611
skos:closeMatch	https://www.freebase.com/m/02p8r

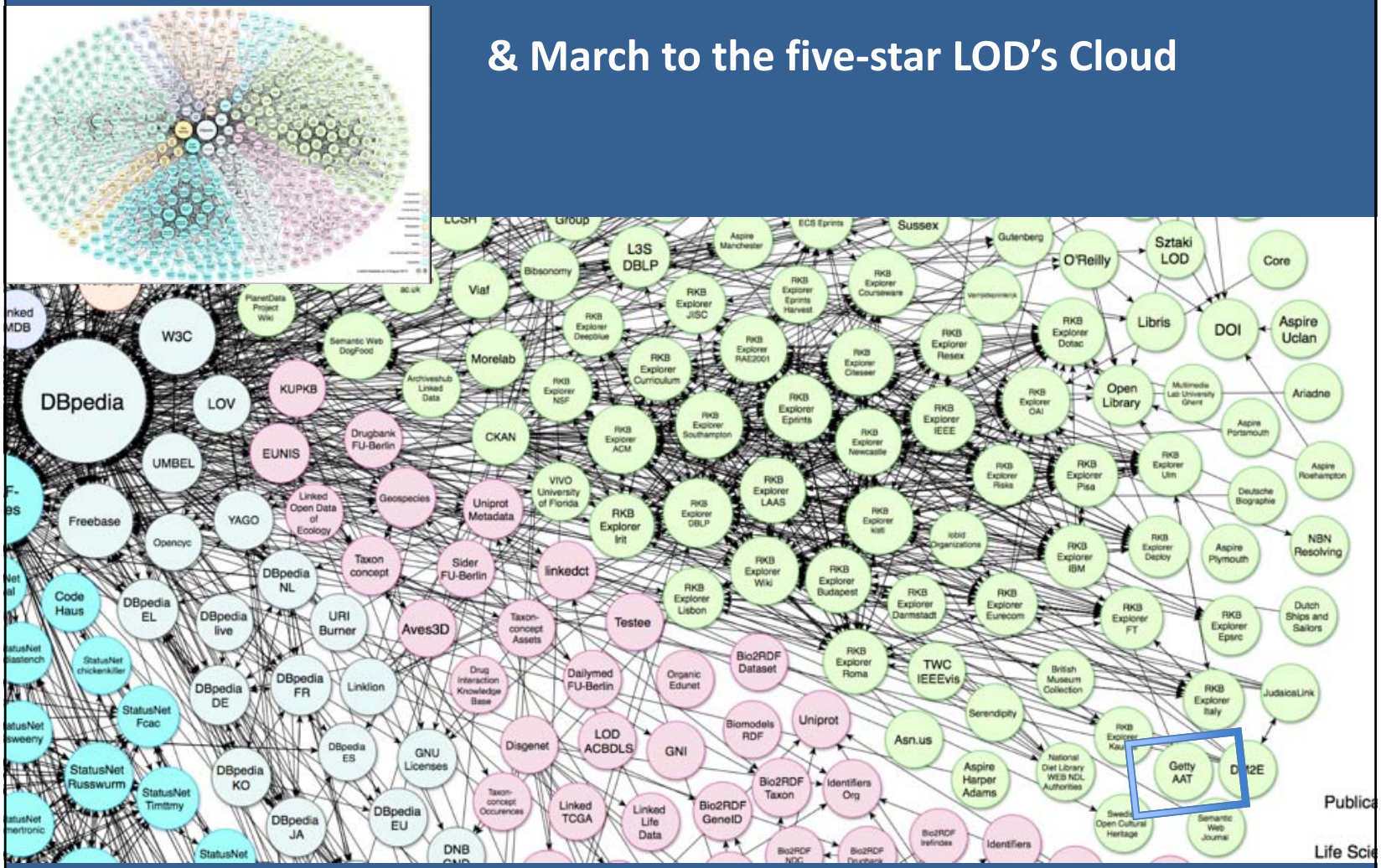
#this (geo:SpatialThing)

dcterms:isPartOf	http://nomisma.org/id/ionia#this
geo:lat	37.939722
geo:long	27.340833

Export

GitHub File RDF/XML RDF/TTL JSON-LD KML

& March to the five-star LOD's Cloud

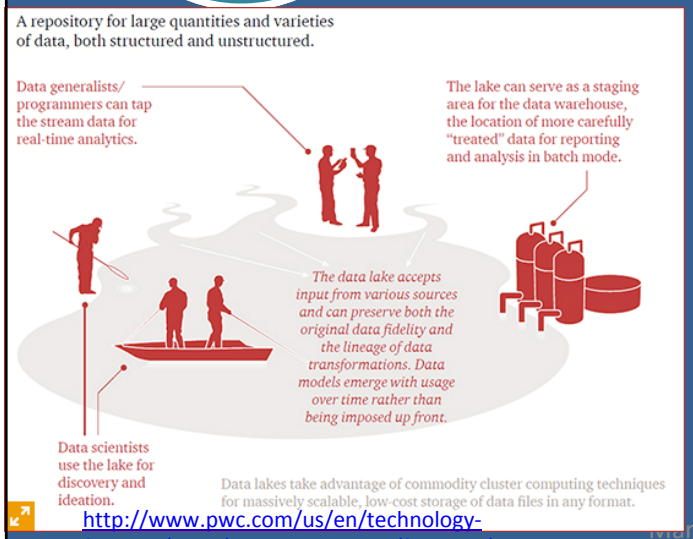


http://lod-cloud.net/ 2014-08

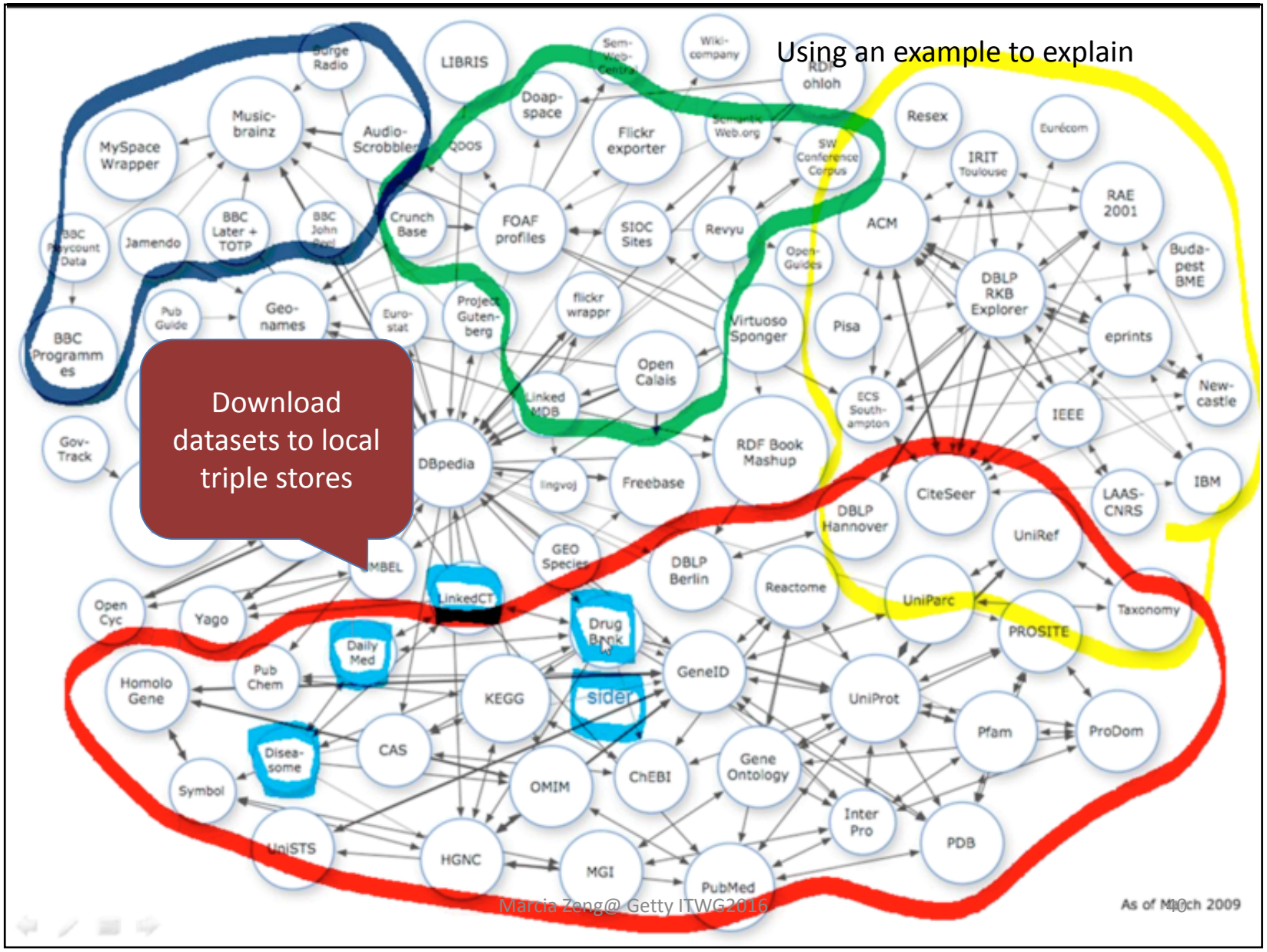
III. For Data Lakes (repositories)

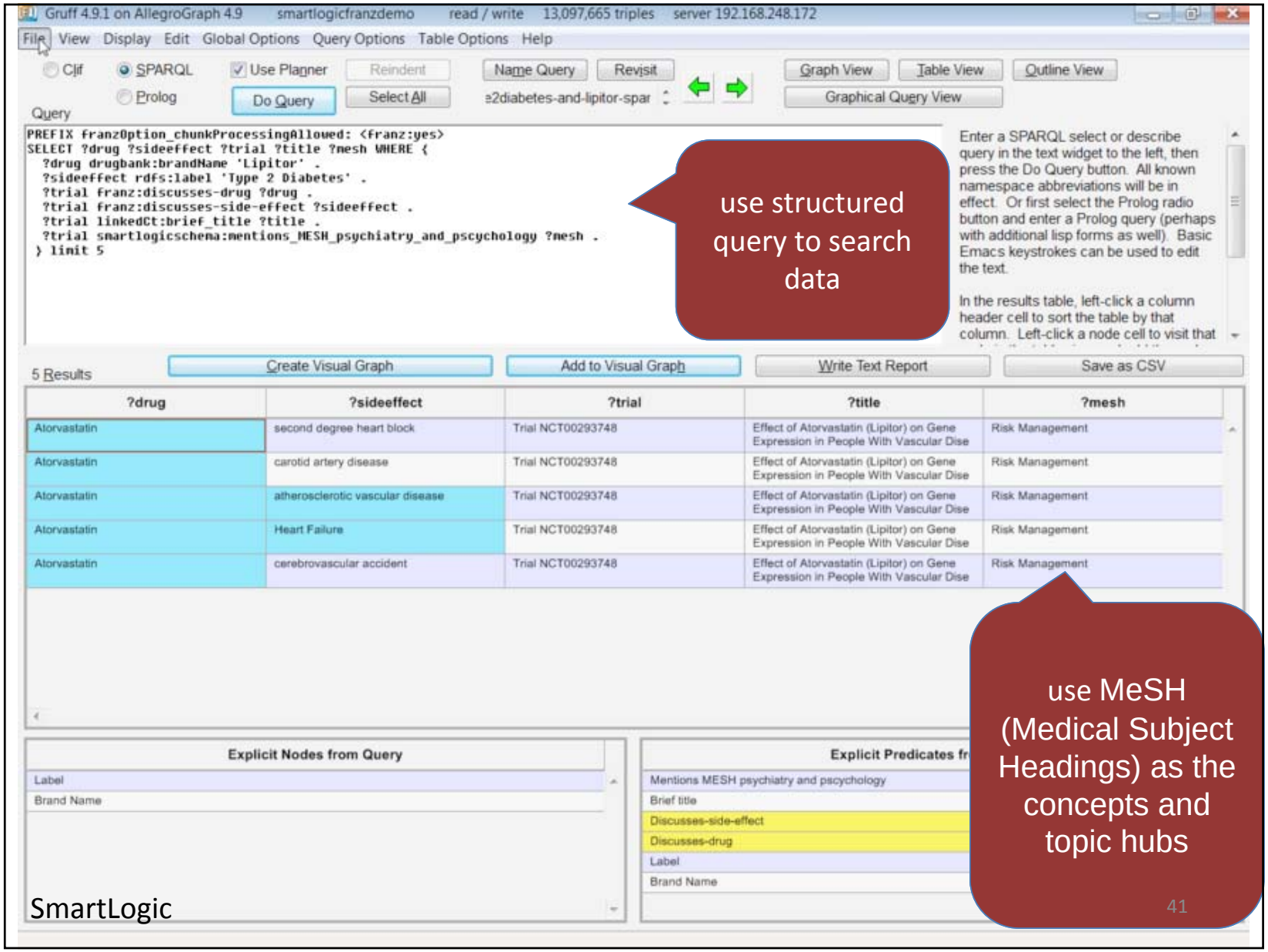
1. Managing the interlinking between datasets
2. Data disambiguation
3. Entity alignment
4. Enabling multilingual and cross lingual discoveries

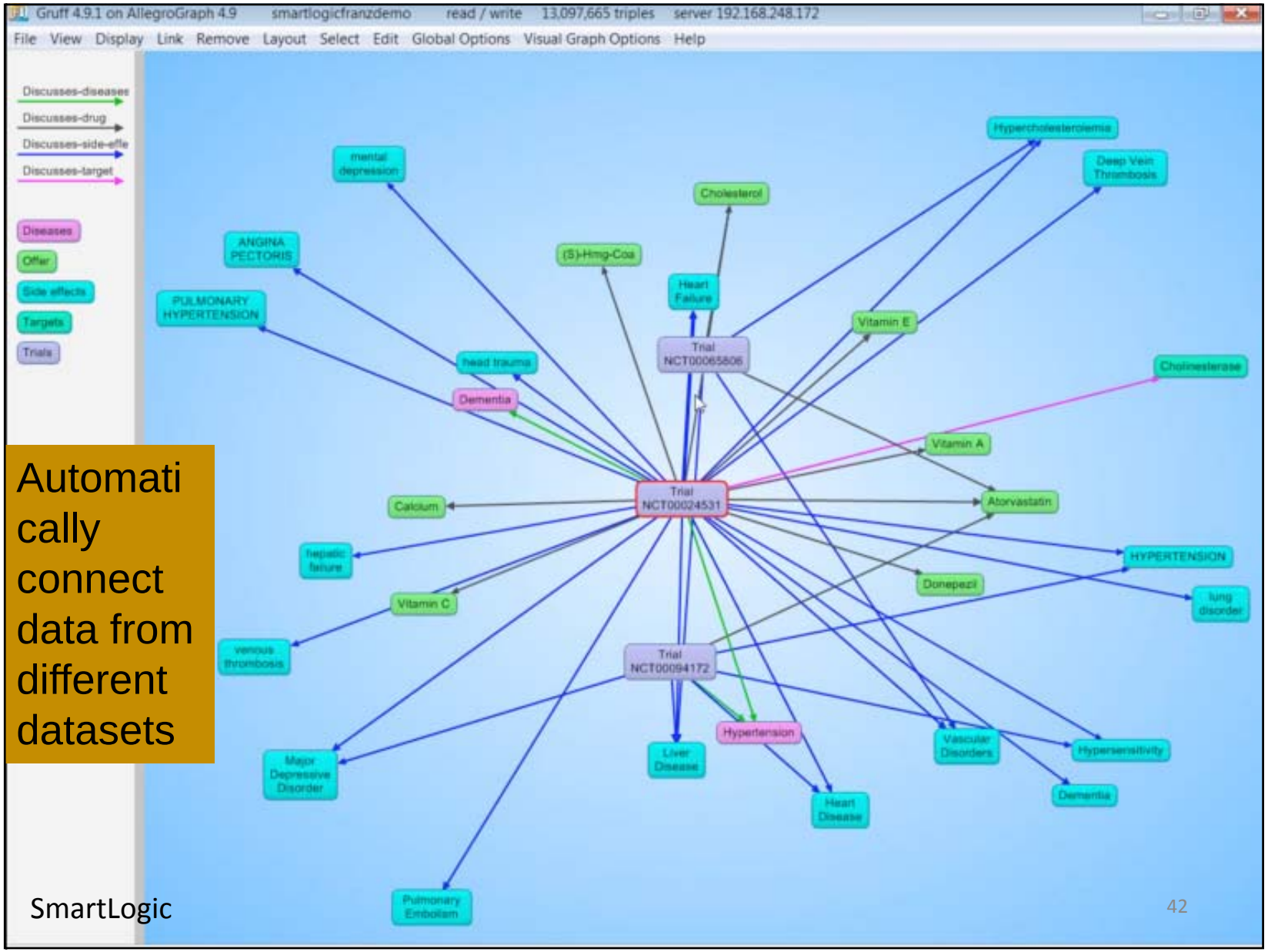
Getty Vocabs



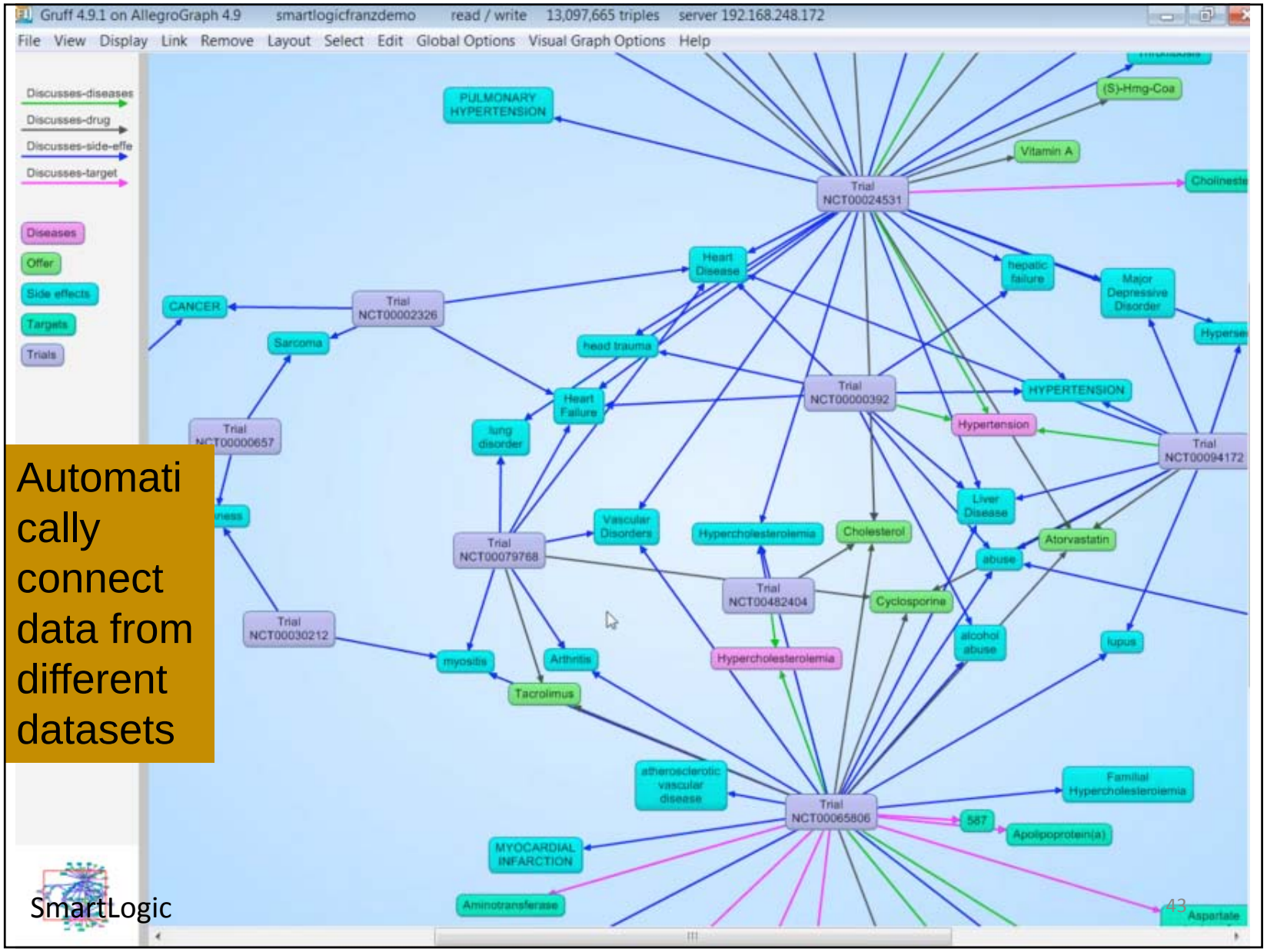
http://www.pwc.com/us/en/technology-forecast/2014/cloud-computing/features/data-lakes.html


Using an example to explain
Download datasets to local triple stores
LIBRIS
Music-brainz
MySpace Wrapper
Audio-Scrobbler
Doap-space
Flickr exporter
FOAF profiles
SIOC Sites
Crunch Base
Jamendo
BBC Later + TOTP
Geo-names
Project Guten-berg
flickr wrappr
Open Calais
Linked MDB
DBpedia
Freebase
RDF Book Mashup
Virtuoso Sponger
ACM
DBLP RKB Explorer
Pisa
eprints
IEEE
New-castle
IBM
LAAS-CNRS
CiteSeer
DBLP Hannover
UniRef
UniParc
Taxonomy
PROSITE
Reactome
DBLP Berlin
GeneID
UniProt
Pfam
ProDom
Gene Ontology
Inter Pro
PDB
PubMed
ChEBI
OMIM
MGI
HGNC
UniSTS
KEGG
CAS
Pub Chem
Homolo Gene
Symbol
Disea-some
Drug Bank
sider
Daily Med
LinkedCT
GEO Species
MBEL
Yago
Open Cyc
Gov-Track
BBC Programmes
Pub Guide
Euro-stat
Revyu
SW Conference Corpus
Resex
IRIT Toulouse
RAE 2001
Buda-pest BME
ECS South-ampton
As of March 2009


Gruff 4.9.1 on AllegroGraph 4.9 smartlogicfranzdemo read / write 13,097,665 triples server 192.168.248.172
File View Display Edit Global Options Query Options Table Options Help
Clif SPARQL Prolog Use Planner Reindent Do Query Select All Name Query Revisit Graph View Table View Outline View Graphical Query View
Query
PREFIX franzOption_chunkProcessingAllowed: <franz:yes>
SELECT ?drug ?sideeffect ?trial ?title ?mesh WHERE {
?drug drugbank:brandName 'Lipitor' .
?sideeffect rdfs:label 'Type 2 Diabetes' .
?trial franz:discusses-drug ?drug .
?trial franz:discusses-side-effect ?sideeffect .
?trial linkedCt:brief_title ?title .
?trial smartlogicschema:mentions_MESH_psychiatry_and_pscychology ?mesh .
} limit 5
use structured query to search data
Enter a SPARQL select or describe query in the text widget to the left, then press the Do Query button. All known namespace abbreviations will be in effect. Or first select the Prolog radio button and enter a Prolog query (perhaps with additional lisp forms as well). Basic Emacs keystrokes can be used to edit the text.
In the results table, left-click a column header cell to sort the table by that column. Left-click a node cell to visit that
5 Results
Create Visual Graph
Add to Visual Graph
Write Text Report
Save as CSV
?drug ?sideeffect ?trial ?title ?mesh
Atorvastatin second degree heart block Trial NCT00293748 Effect of Atorvastatin (Lipitor) on Gene Expression in People With Vascular Dise Risk Management
Atorvastatin carotid artery disease Trial NCT00293748 Effect of Atorvastatin (Lipitor) on Gene Expression in People With Vascular Dise Risk Management
Atorvastatin atherosclerotic vascular disease Trial NCT00293748 Effect of Atorvastatin (Lipitor) on Gene Expression in People With Vascular Dise Risk Management
Atorvastatin Heart Failure Trial NCT00293748 Effect of Atorvastatin (Lipitor) on Gene Expression in People With Vascular Dise Risk Management
Atorvastatin cerebrovascular accident Trial NCT00293748 Effect of Atorvastatin (Lipitor) on Gene Expression in People With Vascular Dise Risk Management
use MeSH (Medical Subject Headings) as the concepts and topic hubs
Explicit Nodes from Query
Label
Brand Name
Explicit Predicates fr
Mentions MESH psychiatry and pscychology
Brief title
Discusses-side-effect
Discusses-drug
Label
Brand Name
SmartLogic
41

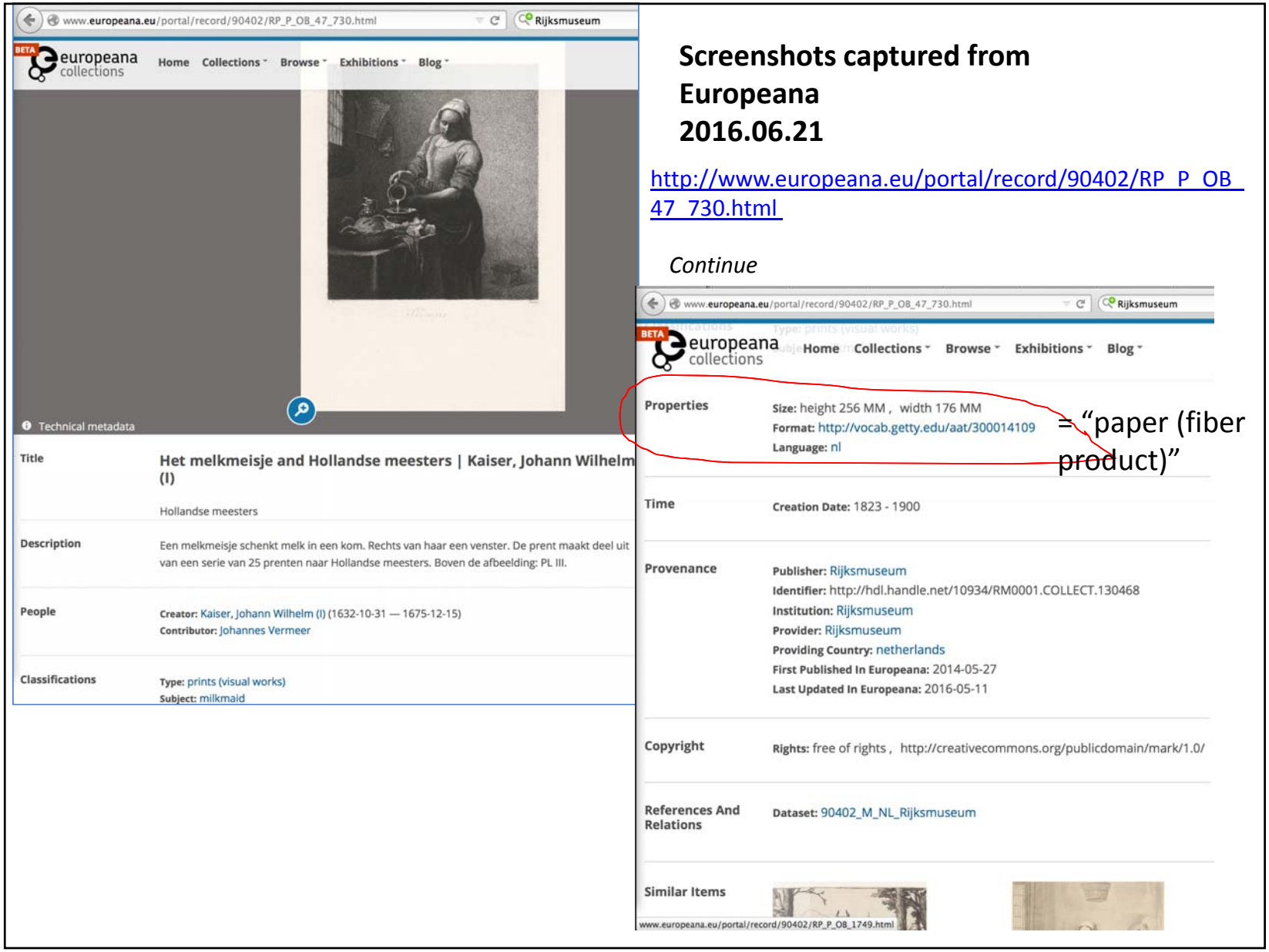

Gruff 4.9.1 on AllegroGraph 4.9
smartlogicfranzdemo
read / write
13,097,665 triples
server 192.168.248.172
File View Display Link Remove Layout Select Edit Global Options Visual Graph Options Help
Discusses-diseases
Discusses-drug
Discusses-side-effe
Discusses-target
Diseases
Offer
Side effects
Targets
Trials
Automatically connect data from different datasets
mental depression
ANGINA PECTORIS
PULMONARY HYPERTENSION
head trauma
Dementia
Cholesterol
(S)-Hmg-Coa
Heart Failure
Trial NCT00065806
Hypercholesterolemia
Deep Vein Thrombosis
Vitamin E
Cholinesterase
Vitamin A
Trial NCT00024531
Calcium
Atorvastatin
hepatic failure
Vitamin C
Donepezil
HYPERTENSION
lung disorder
venous thrombosis
Trial NCT00094172
Hypertension
Liver Disease
Major Depressive Disorder
Vascular Disorders
Hypersensitivity
Dementia
Heart Disease
Pulmonary Embolism

SmartLogic

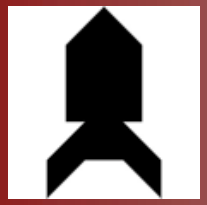


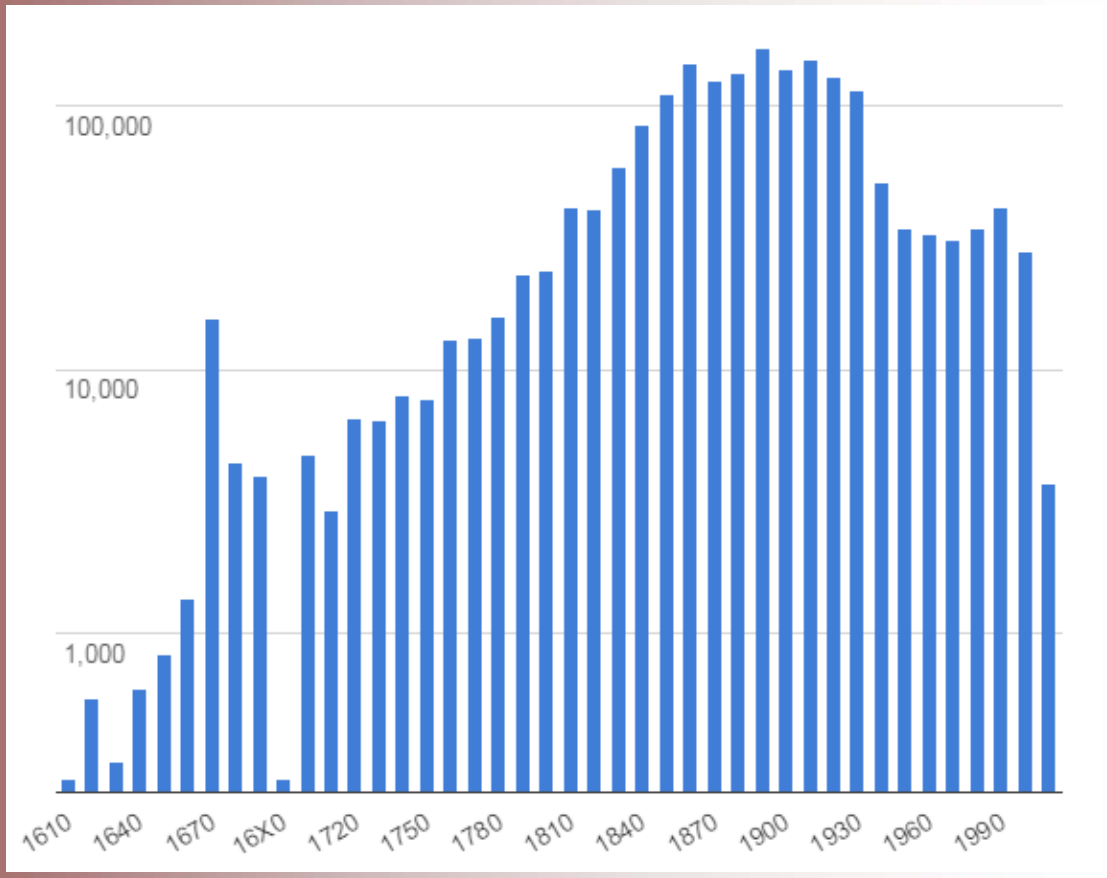
Automatically connect data from different datasets

SmartLogic


www.europeana.eu/portal/record/90402/RP_P_OB_47_730.html
Rijksmuseum
europeana collections
Home Collections Browse Exhibitions Blog
Technical metadata
Title
Het melkmeisje and Hollandse meesters | Kaiser, Johann Wilhelm (I)
Hollandse meesters
Description
Een melkmeisje schenkt melk in een kom. Rechts van haar een venster. De prent maakt deel uit van een serie van 25 prenten naar Hollandse meesters. Boven de afbeelding: PL III.
People
Creator: Kaiser, Johann Wilhelm (I) (1632-10-31 — 1675-12-15)
Contributor: Johannes Vermeer
Classifications
Type: prints (visual works)
Subject: milkmaid
Screenshots captured from Europeana
2016.06.21
http://www.europeana.eu/portal/record/90402/RP_P_OB_47_730.html
Continue
Properties
Size: height 256 MM , width 176 MM
Format: http://vocab.getty.edu/aat/300014109
Language: nl
= "paper (fiber product)"
Time
Creation Date: 1823 - 1900
Provenance
Publisher: Rijksmuseum
Identifier: http://hdl.handle.net/10934/RM0001.COLLECT.130468
Institution: Rijksmuseum
Provider: Rijksmuseum
Providing Country: netherlands
First Published In Europeana: 2014-05-27
Last Updated In Europeana: 2016-05-11
Copyright
Rights: free of rights , http://creativecommons.org/publicdomain/mark/1.0/
References And Relations
Dataset: 90402_M_NL_Rijksmuseum
Similar Items
www.europeana.eu/portal/record/90402/RP_P_OB_1749.html

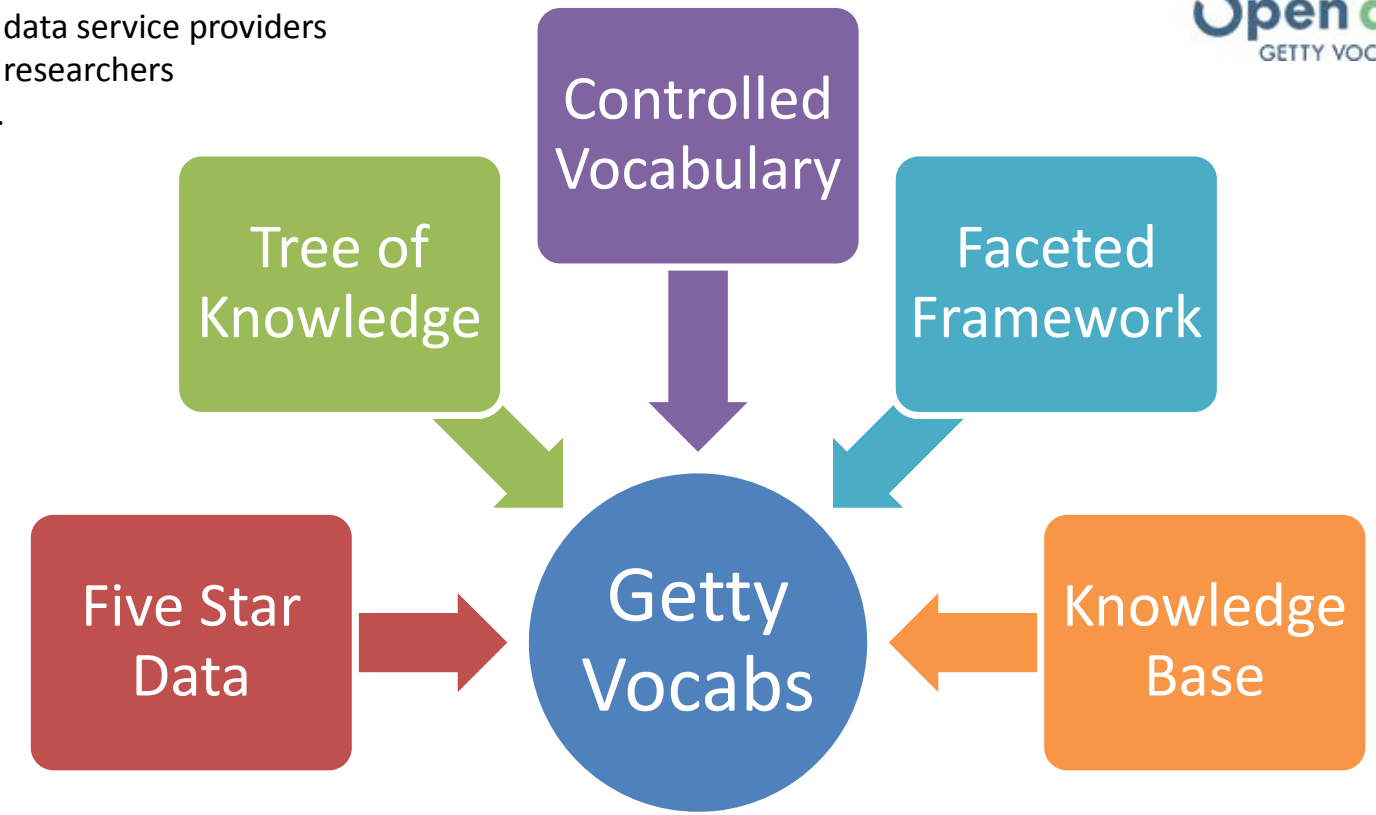
Search millions of historic newspapers on Europeana using a simple query like {skos_concept:"http://vocab.getty.edu/aat/300026656"}





Conclusions

Why does any dataset need to care about the "Getty Vocabularies"?"

If any of your needs can be met by applying the Getty vocabularies, then ride on it to reach the five-star level!

- ✓ For vocab creators/managers
- ✓ For LOD data creators
- ✓ For data service providers
- ✓ For researchers
- ✓ … …