INFO/CS 4302
Web Information Systems

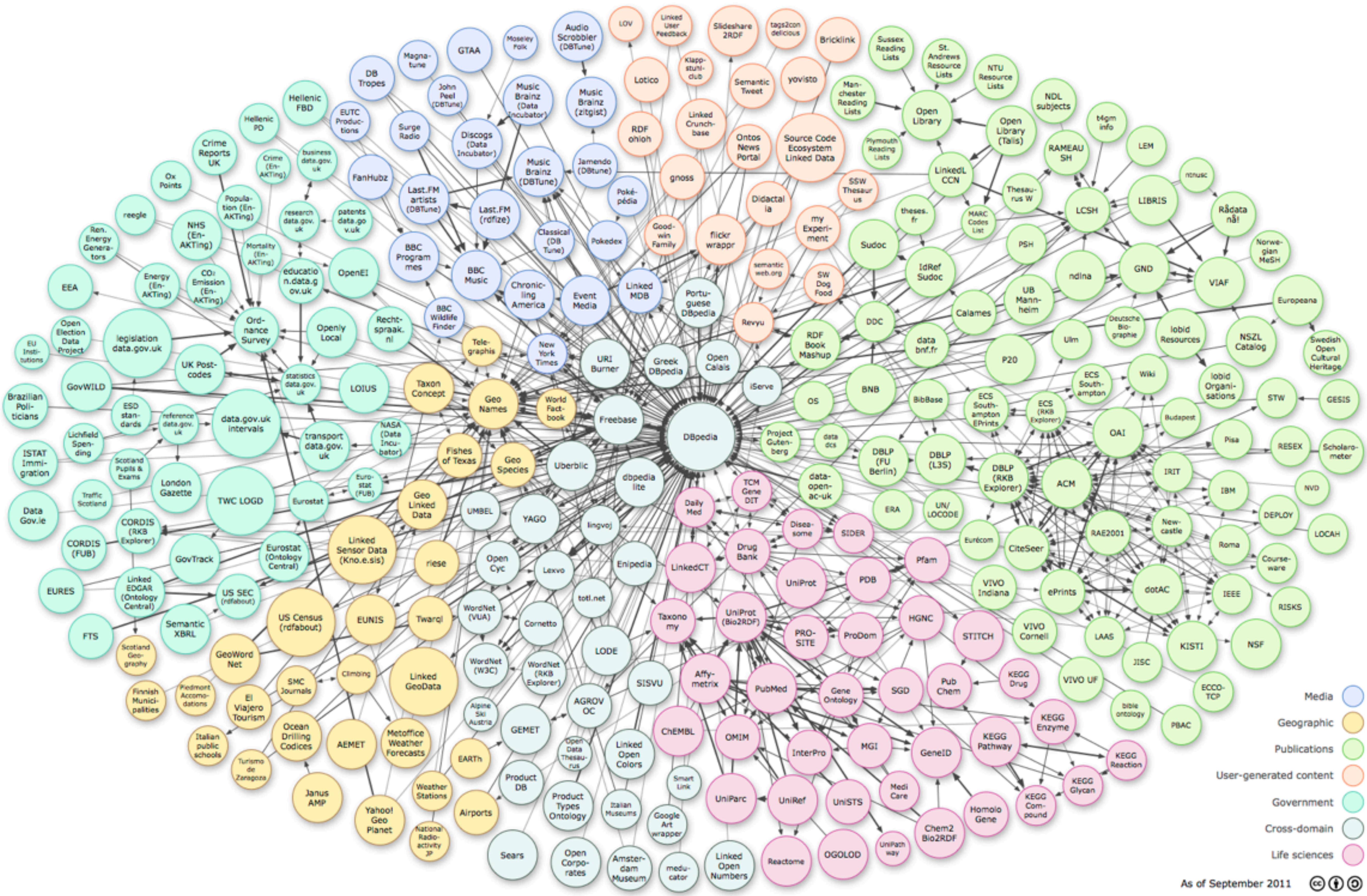
FT 2012
Lecture 17: SKOS, SPARQL

- Bernhard Haslhofer -

Plan for today...

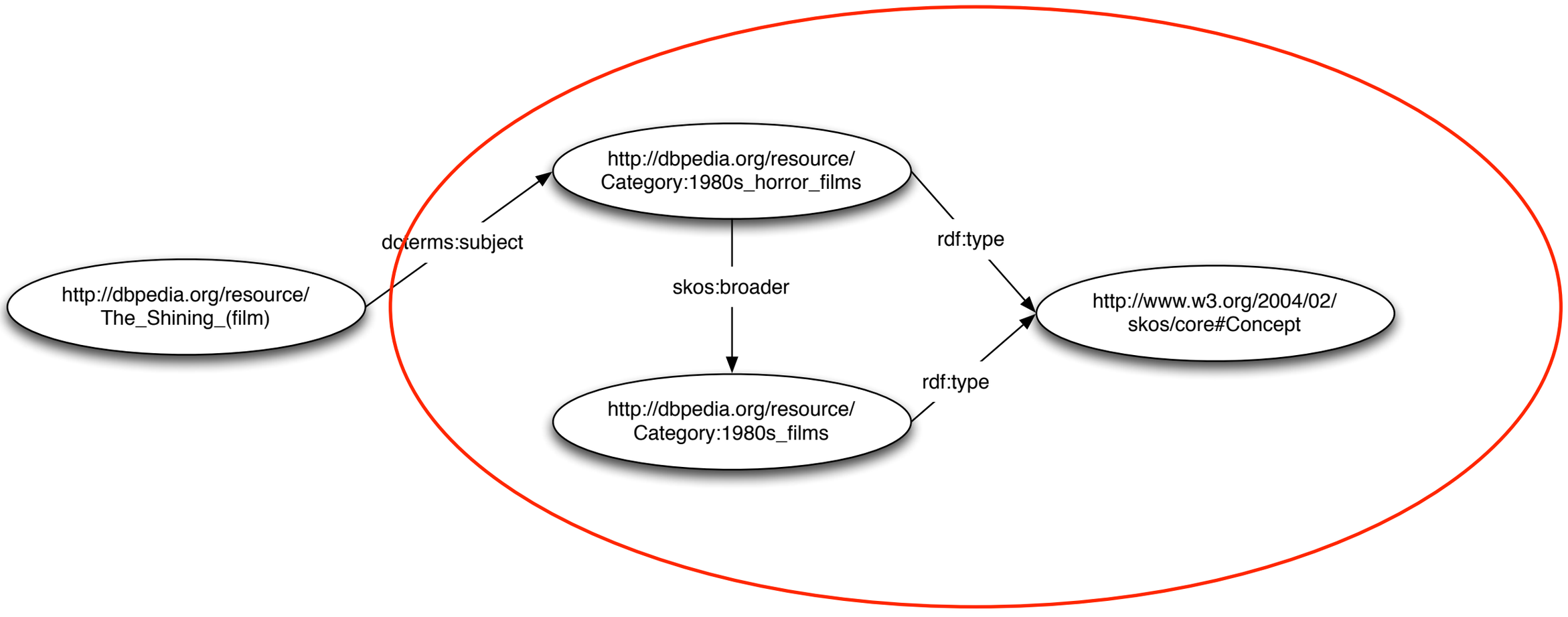
- Context & Recap
- SKOS
- SPARQL
- Questions

CONTEXT & RECAP


LOV
Linked User Feedback
Slideshare 2RDF
tags2con delicious
Bricklink
Moseley Folk
Audio Scrobbler (DBTune)
GTAA
Magna-tune
DB Tropes
John Peel (DBTune)
Music Brainz (Data Incubator)
Music Brainz (zitgist)
Lotico
Klapp-stuhl-club
Semantic Tweet
yovisto
Sussex Reading Lists
St. Andrews Resource Lists
NTU Resource Lists
Man-chester Reading Lists
Open Library
NDL subjects
Hellenic FBD
Hellenic PD
EUTC Produc-tions
Surge Radio
Discogs (Data Incubator)
RDF ohloh
Linked Crunch-base
Ontos News Portal
Source Code Ecosystem Linked Data
Plymouth Reading Lists
Open Library (Talis)
t4gm info
RAMEAU SH
LEM
Crime Reports UK
Crime (En-AKTing)
business data.gov.uk
FanHubz
Last.FM artists (DBTune)
Music Brainz (DBTune)
Jamendo (DBtune)
gnoss
SSW Thesaur us
LinkedL CCN
Thesau-rus W
ntnusc
Ox Points
Popula-tion (En-AKTing)
Poké-pédia
Didactal ia
theses. fr
MARC Codes List
LCSH
LIBRIS
Rådata nå!
reegle
NHS (En-AKTing)
research data.gov. uk
patents data.go v.uk
Last.FM (rdfize)
Classical (DB Tune)
Pokedex
Good-win Family
flickr wrappr
my Experi-ment
Sudoc
PSH
Norwe-gian MeSH
Ren. Energy Genera-tors
Mortality (En-AKTing)
BBC Program mes
semantic web.org
IdRef Sudoc
EEA
Energy (En-AKTing)
CO2 Emission (En-AKTing)
educatio n.data.g ov.uk
OpenEI
BBC Music
Chronic-ling America
Event Media
Linked MDB
Portu-guese DBpedia
SW Dog Food
ndlna
GND
VIAF
Europeana
Open Election Data Project
legislation data.gov.uk
Ord-nance Survey
Openly Local
Recht-spraak. nl
BBC Wildlife Finder
Tele-graphis
Revyu
DDC
Calames
UB Mann-heim
Deutsche Bio-graphie
lobid Resources
NSZL Catalog
Swedish Open Cultural Heritage
EU Insti-tutions
UK Post-codes
statistics data.gov. uk
New York Times
URI Burner
Greek DBpedia
Open Calais
RDF Book Mashup
data bnf.fr
P20
Ulm
ECS South-ampton
Wiki
lobid Organi-sations
Brazilian Poli-ticians
GovWILD
ESD stan-dards
reference data.gov. uk
data.gov.uk intervals
LOIUS
Taxon Concept
Geo Names
World Fact-book
Freebase
DBpedia
iServe
OS
BNB
BibBase
ECS South-ampton EPrints
ECS (RKB Explorer)
OAI
Budapest
STW
GESIS
Lichfield Spen-ding
transport data.gov. uk
NASA (Data Incu-bator)
Fishes of Texas
Geo Species
Uberblic
dbpedia lite
Project Guten-berg
data dcs
DBLP (FU Berlin)
DBLP (L3S)
DBLP (RKB Explorer)
ACM
Pisa
RESEX
Scholaro-meter
ISTAT Immi-gration
Scotland Pupils & Exams
London Gazette
TWC LOGD
Eurostat
Euro-stat (FUB)
Geo Linked Data
UMBEL
YAGO
TCM Gene DIT
data-open-ac-uk
ERA
UN/ LOCODE
IRIT
IBM
NVD
Data Gov.ie
Traffic Scotland
CORDIS (RKB Explorer)
Daily Med
lingvoj
Disea-some
SIDER
Eurécom
CiteSeer
RAE2001
New-castle
DEPLOY
CORDIS (FUB)
GovTrack
Eurostat (Ontology Central)
Linked Sensor Data (Kno.e.sis)
riese
Open Cyc
Lexvo
Enipedia
LinkedCT
Drug Bank
Pfam
Roma
Course-ware
LOCAH
EURES
Linked EDGAR (Ontology Central)
US SEC (rdfabout)
totl.net
UniProt
PDB
VIVO Indiana
ePrints
dotAC
IEEE
RISKS
FTS
Semantic XBRL
US Census (rdfabout)
EUNIS
Twarql
WordNet (VUA)
Cornetto
LODE
Taxono my
UniProt (Bio2RDF)
PRO-SITE
ProDom
HGNC
STITCH
VIVO Cornell
LAAS
KISTI
NSF
Scotland Geo-graphy
GeoWord Net
SMC Journals
Climbing
Linked GeoData
WordNet (W3C)
WordNet (RKB Explorer)
AGROV OC
SISVU
Affy-metrix
PubMed
Gene Ontology
SGD
Pub Chem
KEGG Drug
JISC
VIVO UF
bible ontology
ECCO-TCP
Finnish Munici-palities
Piedmont Accomo-dations
El Viajero Tourism
Ocean Drilling Codices
AEMET
Metoffice Weather Forecasts
Alpine Ski Austria
GEMET
ChEMBL
OMIM
MGI
GeneID
KEGG Pathway
KEGG Enzyme
PBAC
Italian public schools
Turismo de Zaragoza
EARTh
Open Data Thesau-rus
Linked Open Colors
InterPro
KEGG Reaction
Janus AMP
Yahoo! Geo Planet
Weather Stations
Airports
Product DB
Product Types Ontology
Italian Museums
Smart Link
UniParc
UniRef
UniSTS
Medi Care
KEGG Com-pound
KEGG Glycan
National Radio-activity JP
Sears
Open Corpo-rates
Amster-dam Museum
Google Art wrapper
medu-cator
Linked Open Numbers
Reactome
OGOLOD
UniPath way
Chem2 Bio2RDF
Homolo Gene
Media
Geographic
Publications
User-generated content
Government
Cross-domain
Life sciences
As of September 2011

SPARQL

- A language for describing knowledge organization systems (taxonomies, thesauri, classification schemes)



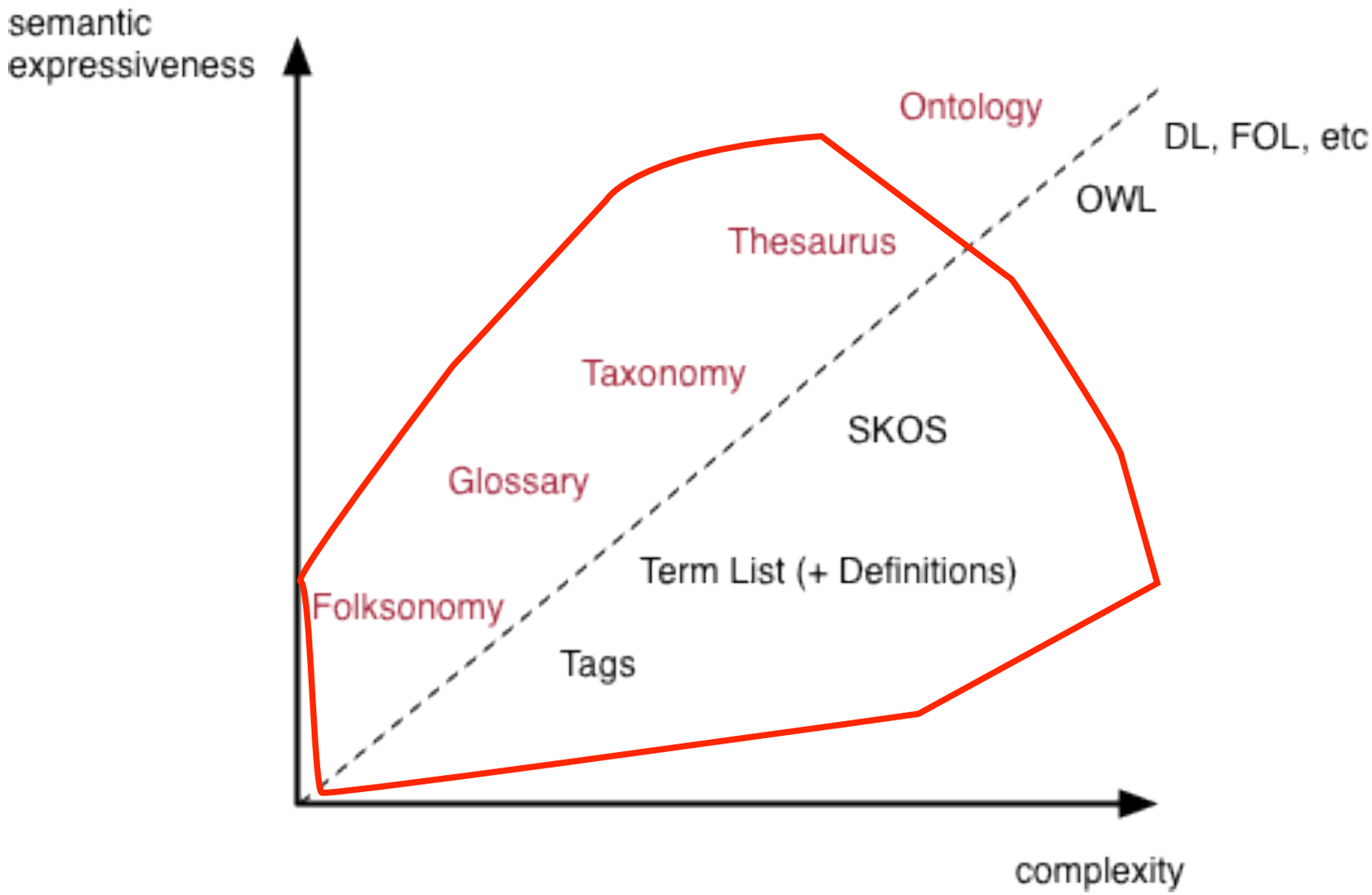
SPARQL

- A query language and protocol for accessing RDF data on the Web

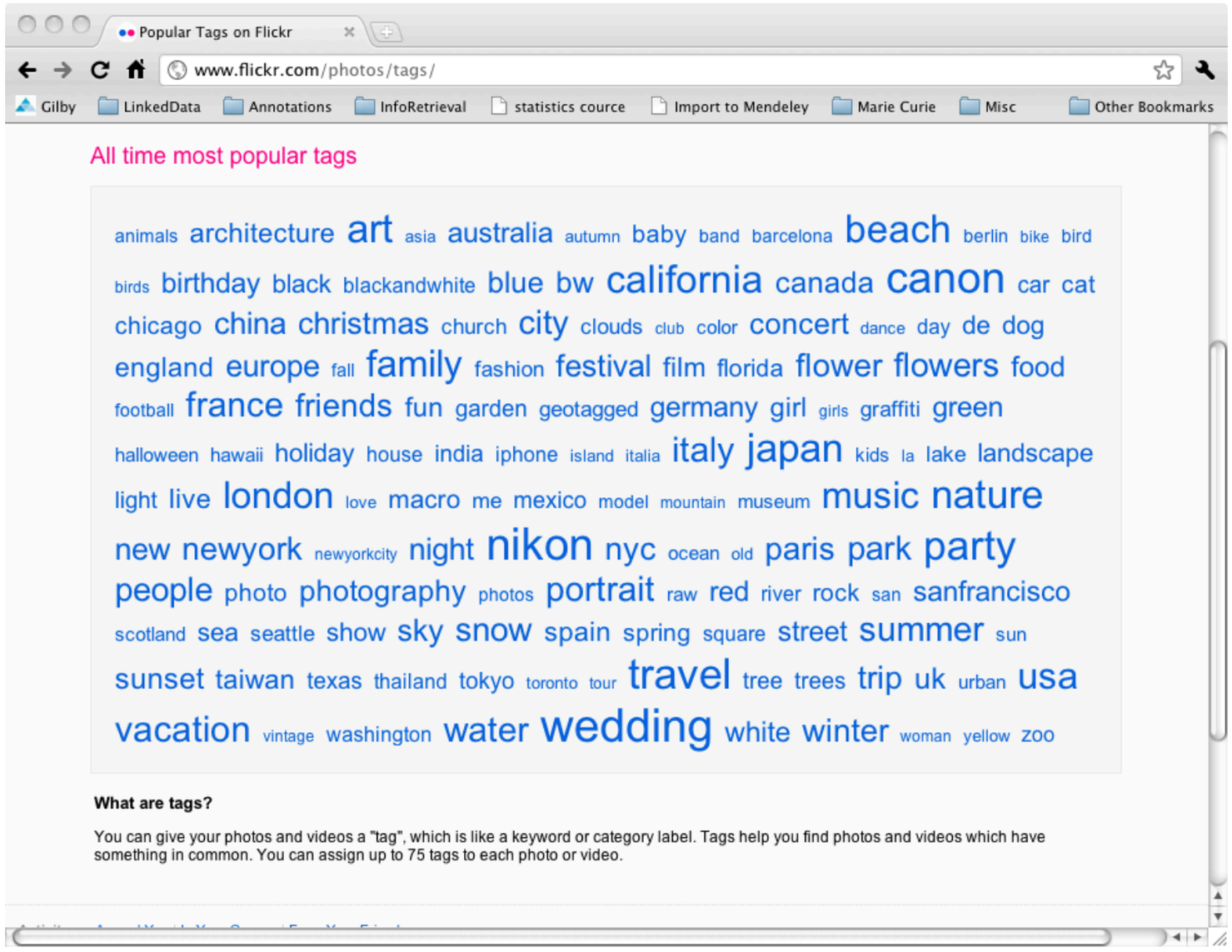
```
SELECT DISTINCT ?x
WHERE {
      ?x dcterms:subject
      <http://dbpedia.org/resource/Category:1980s_horror_films> .
}
```

SKOS

Knowledge Organization Systems (KOS)



Folksonomy Examples



twitter

#info4302

North American Product Classification System

Site Resources

- What's New

Contact Us

Email Us

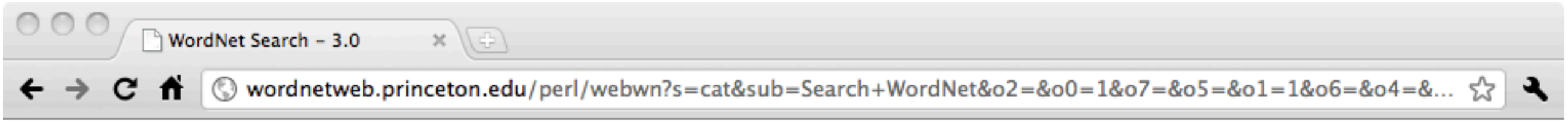
Introduction

In February 1999, the statistical agencies of Canada, Mexico, and the United States launched a joint multi-phase initiative to develop a comprehensive demand-oriented product classification, known as the North American Product Classification System (NAPCS). Work to date has focused on the products produced by service industries in 12 NAICS sectors 48-49 through 81. With that work nearing completion, this web page provides an overview of and progress report on the NAPCS initiative and presents the final versions of the product lists developed so far for the service industries included in those 12 sectors.

Overview and Progress Report

Final NAPCS Product Lists

- Description of product lists
- View PDF files of individual product lists
- Download sequential Excel file of all product lists



WordNet Search - 3.0 - WordNet home page - Glossary - Help

Word to search for: cat Search WordNet

Display Options: (Select option to change) Change

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

Noun

- S: (n) **cat**, true cat (feline mammal usually having thick soft fur and no ability to roar: domestic cats; wildcats)
- S: (n) guy, **cat**, hombre, bozo (an informal term for a youth or man) *"a nice guy"; "the guy's only doing it for some doll"*
- S: (n) **cat** (a spiteful woman gossip) *"what a cat she is!"*
- S: (n) kat, khat, qat, quat, **cat**, Arabian tea, African tea (the leaves of the shrub Catha edulis which are chewed like tobacco or used to make tea; has the effect of a euphoric stimulant) *"in Yemen kat is used daily by 85% of adults"*
- S: (n) cat-o'-nine-tails, **cat** (a whip with nine knotted cords) *"British sailors feared the cat"*
- S: (n) Caterpillar, **cat** (a large tracked vehicle that is propelled by two endless metal belts; frequently used for moving earth in construction and farm work)
- S: (n) big cat, **cat** (any of several large cats typically able to roar and living in the wild)
- S: (n) computerized tomography, computed tomography, CT, computerized axial tomography, computed axial tomography, **CAT** (a method of examining body organs by scanning them with X rays and using a computer to construct a series of cross-sectional scans along a single axis)

Verb

- S: (v) **cat** (beat with a cat-o'-nine-tails)
- S: (v) vomit, vomit up, purge, cast, sick, **cat**, be sick, disgorge, regorge, retch, puke, barf, spew, spue, chuck, upchuck, honk, regurgitate, throw up (eject the contents of the stomach through the mouth) *"After drinking too much, the students vomited"; "He purged continuously"; "The patient regurgitated the food we gave him last night"*

WordNet home page

craigslist: ithaca classifieds f

ithaca.craigslist.org

craigslist

post to classifieds
my account
help, faq, abuse, legal

search craigslist

for sale >

event calendar

S	M	T	W	T	F	S
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5
6	7	8	9	10	11	12

avoid scams & fraud
personal safety tips
craigslist blog
craigslist factsheet
best-of-craigslist
craigslist TV
craigslist movie & dvd
craigslist foundation
craigconnects

system status

terms of use about
privacy help

ithaca w

community

activities
artists
childcare
general
groups
pets
events
lost+found
musicians
local news
politics
rideshare
volunteers
classes

personals

strictly platonic
women seek women
women seeking men
men seeking women
men seeking men
misc romance
casual encounters
missed connections
rants and raves

discussion forums

1099
apple
arts
atheist
autos
beauty
bikes
celebs
comp
crafts
diet
divorce
dying
eco
educ
etiquet
feedbk
film
fitness
gifts
haiku
health
help
history
housing
jobs
jokes
kink
l.t.r.
legal
linux
loc pol
m4m
money
motocy
music
npo
open
pets
philos
politic
psych
queer
recover
religion
rofo
science
shop
spirit
sports
t.v.
tax
testing
transg
travel
vegan
w4w

housing

apts / housing
rooms / shared
sublets / temporary
housing wanted
housing swap
vacation rentals
parking / storage
office / commercial
real estate for sale

for sale

appliances
antiques
barter
bikes
boats
books
business
computer
free
furniture
general
jewelry
materials
rvs
sporting
tickets
tools
wanted
arts+crafts
auto parts
baby+kids
beauty+hlth
cars+trucks
cds/dvd/vhs
cell phones
clothes+acc
collectibles
electronics
farm+garden
garage sale
household
motorcycles
music instr
photo+video
toys+games
video gaming

services

beauty
creative
computer
cycle
event
financial
legal
automotive
farm+garden
household
labor/move
skill'd trade
real estate
sm biz ads

jobs

accounting+finance
admin / office
arch / engineering
art / media / design
biotech / science
business / mgmt
customer service
education
food / bev / hosp
general labor
government
human resources
internet engineers
legal / paralegal
manufacturing
marketing / pr / ad
medical / health
nonprofit sector
real estate
retail / wholesale
sales / biz dev
salon / spa / fitness
security
skilled trade / craft
software / qa / dba
systems / network
technical support
transport
tv / film / video
web / info design
writing / editing
[ETC]
[part-time]

gigs

crew
event
labor
computer
creative
domestic

nearby cl

albany
allentown
belleville
binghamton
buffalo
catskills
elmira
finger lakes
ithaca
kingston
oneonta
poconos
rochester
scranton
state college
syracuse
twin tiers
utica
watertown
williamsport

us cities
us states
canada
cl worldwide

Amazon.com: Online Shoppi

www.amazon.com

AnnotateIt CALIXTO ON CHROM

amazon

Your Amazon.com | Today's Deals | Gift Cards | Help

The All-New kindle fire HD

Shop by Department | Search | All | Go | Hello. Sign in Your Account | 0 Cart | Wish List

- Unlimited Instant Videos
- MP3s & Cloud Player
  20 million songs, play anywhere
- Amazon Cloud Drive
  5 GB of free storage
- Kindle
- Appstore for Android
  Get Call of Atlantis free today
- Digital Games & Software
- Audible Audiobooks

- Books
- Movies, Music & Games
- **Electronics & Computers**
- Home, Garden & Tools
- Grocery, Health & Beauty
- Toys, Kids & Baby
- Clothing, Shoes & Jewelry
- Sports & Outdoors
- Automotive & Industrial

› Full Store Directory

Electronics

- TV & Video
- Home Audio & Theater
- Camera, Photo & Video
- Cell Phones & Accessories
- Video Games
- MP3 Players & Accessories
- Car Electronics & GPS
- Appliances
- Musical Instruments

Computers

- Laptops, Tablets & Netbooks
- Desktops & Servers
- Computer Accessories & Peripherals
  External drives, mice, networking & more
- Computer Parts & Components
- Software
- PC Games
- Printers & Ink
- Office & School Supplies

Car Electronics Resource Center

›Learn more

Get Charged with Powerful and Versatile Equipment

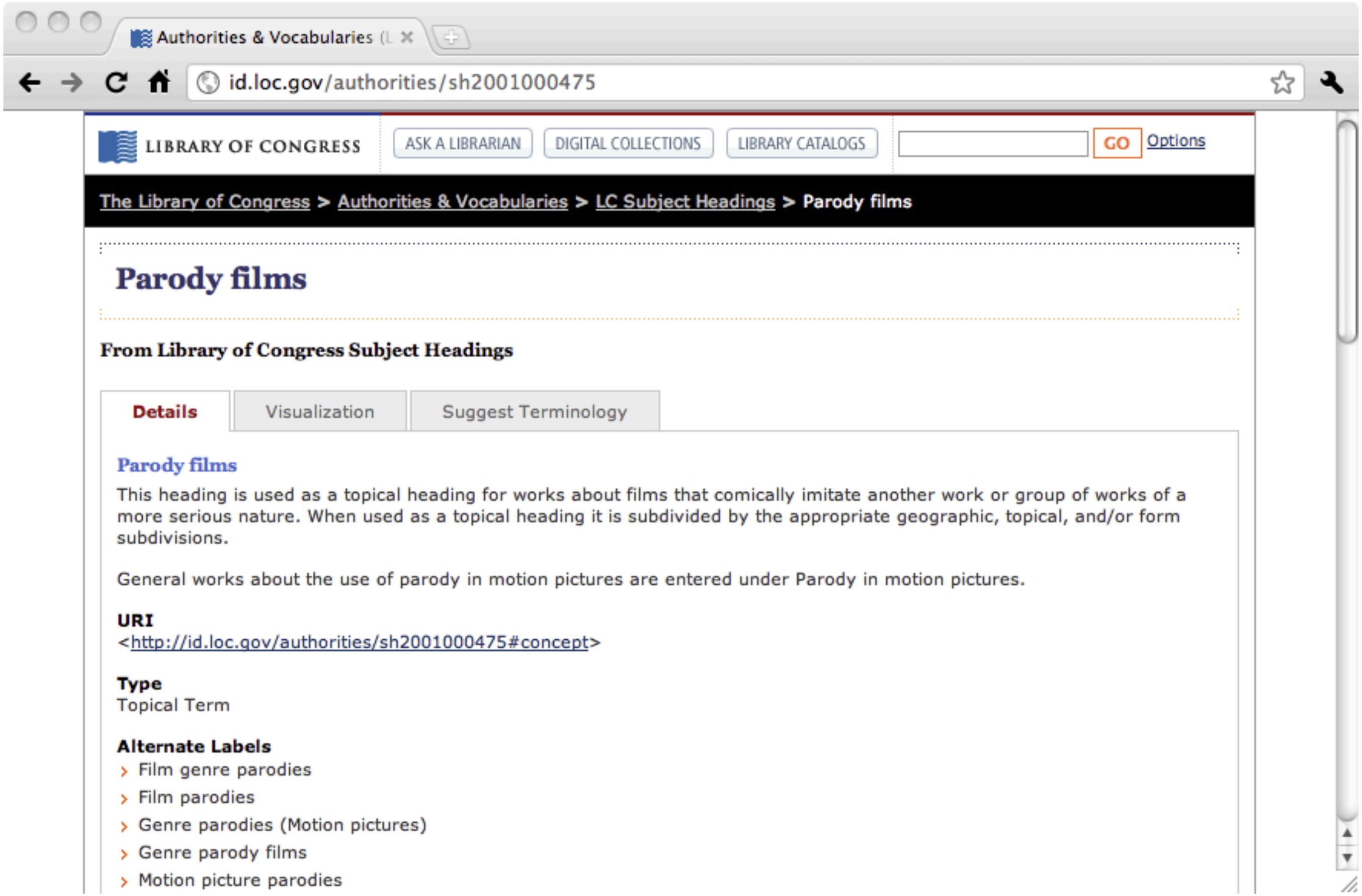
What is SKOS?

- A model for expressing the basic structure and content of concept schemes such as thesauri, classification schemes, taxonomies, folksonomies, and other similar types of controlled vocabularies
- Allows concepts to be composed and published as Linked Data on the Web
- Hides the complexity of OWL - easy to use

SKOS Concepts are...

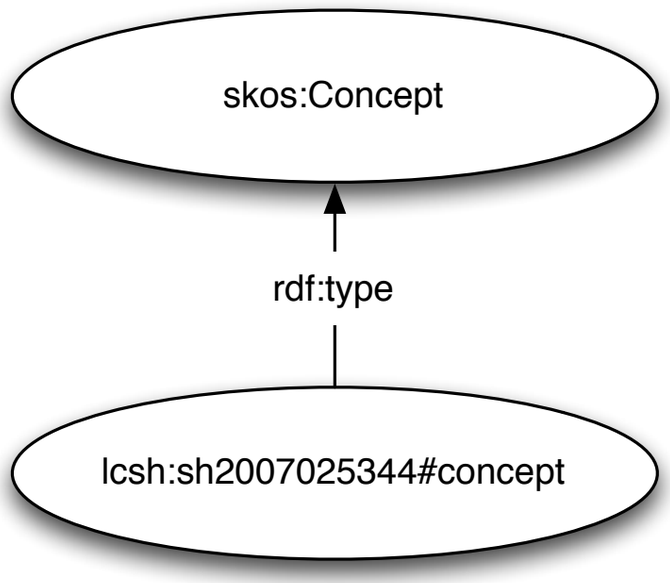
- ... identified by URIs
- ... labeled with 1..* natural language strings
- ... documented with various types of notes
- ... semantically linked to each other
- ... aggregated into concept schemes

Example SKOS Concept



skos:Concept

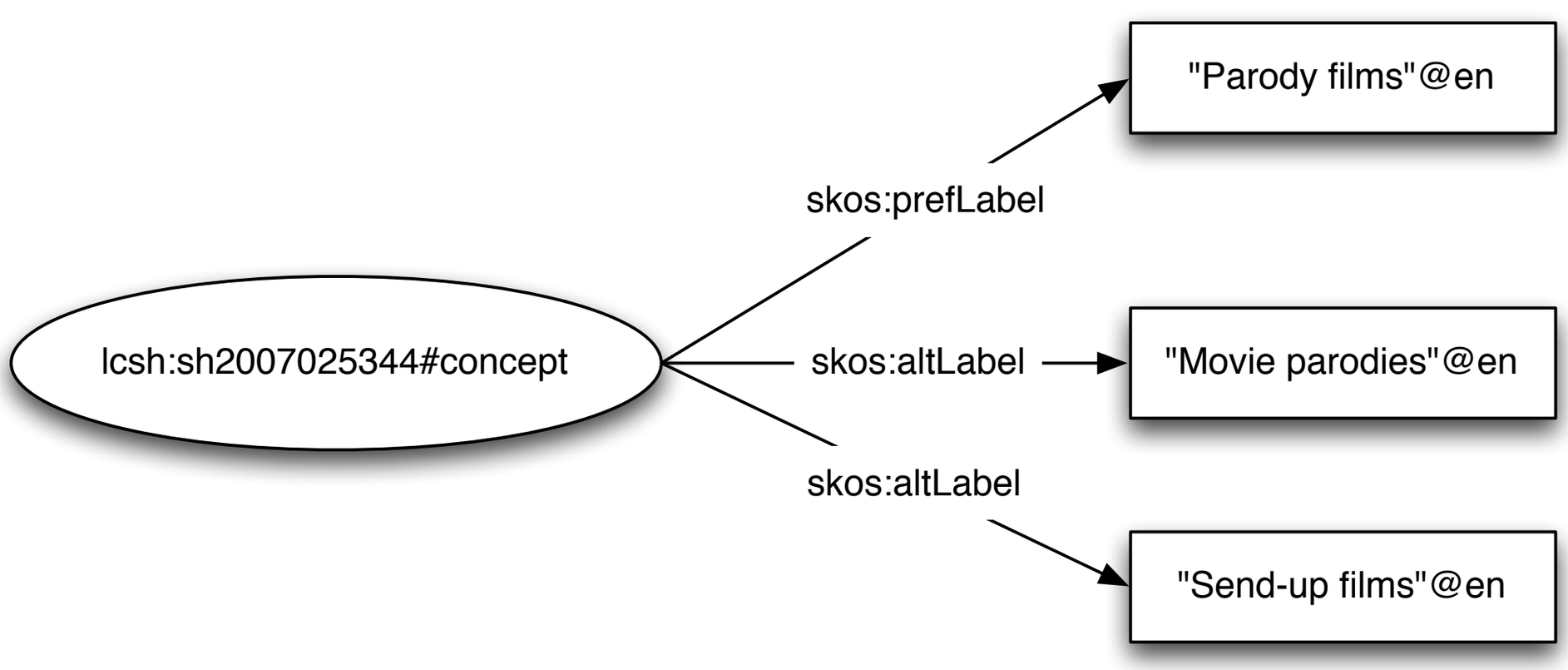
- Concepts are
  - the units of thought: ideas, meanings, categories of objects, etc.
  - abstract entities which are independent of the terms used to label them




```
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix lcsh: <http://id.loc.gov/authorities/> .
```

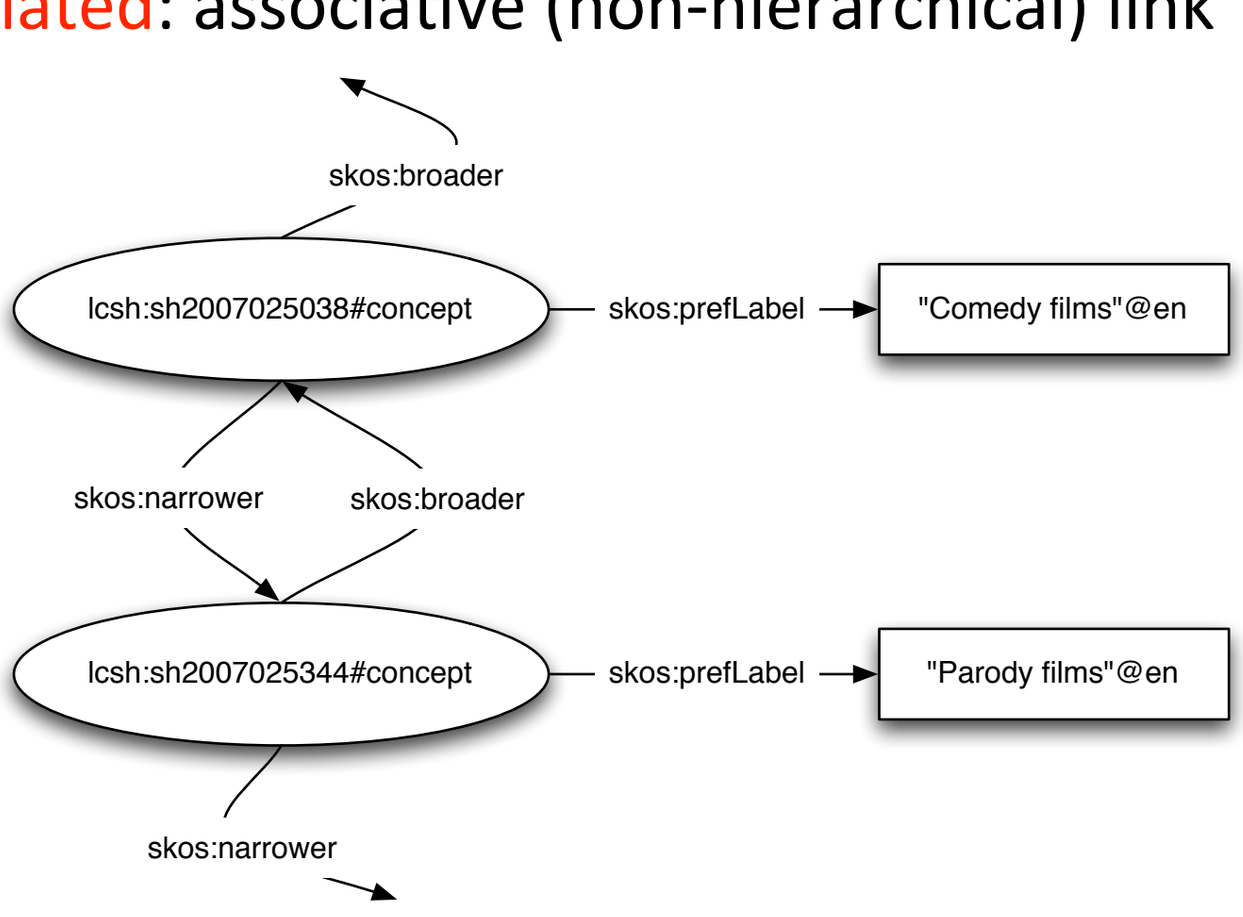
skos:(pref|alt|hidden)Label

- Labels refer to concepts' natural language(s)
  - skos:prefLabel: the preferred lexical label
  - skos:altLabel: alternative lexical labels (e.g., synonyms)
  - skos:hiddenLabel: labels useful for indexing (e.g., typos)



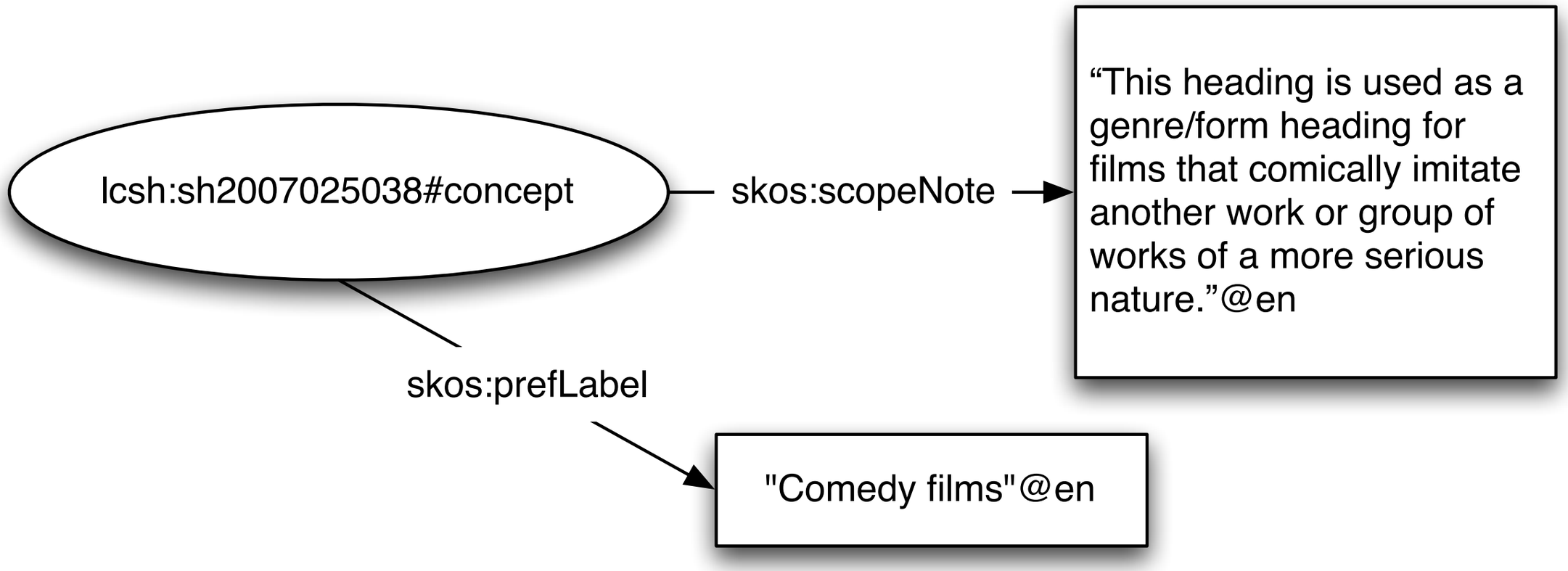
SKOS Semantic Relationships

- The meaning of a concept is also defined by its links to other concepts
  - skos:broader: hierarchical link to a more general concept
  - skos:narrower: hierarchical link to a more specific concept
  - skos:related: associative (non-hierarchical) link



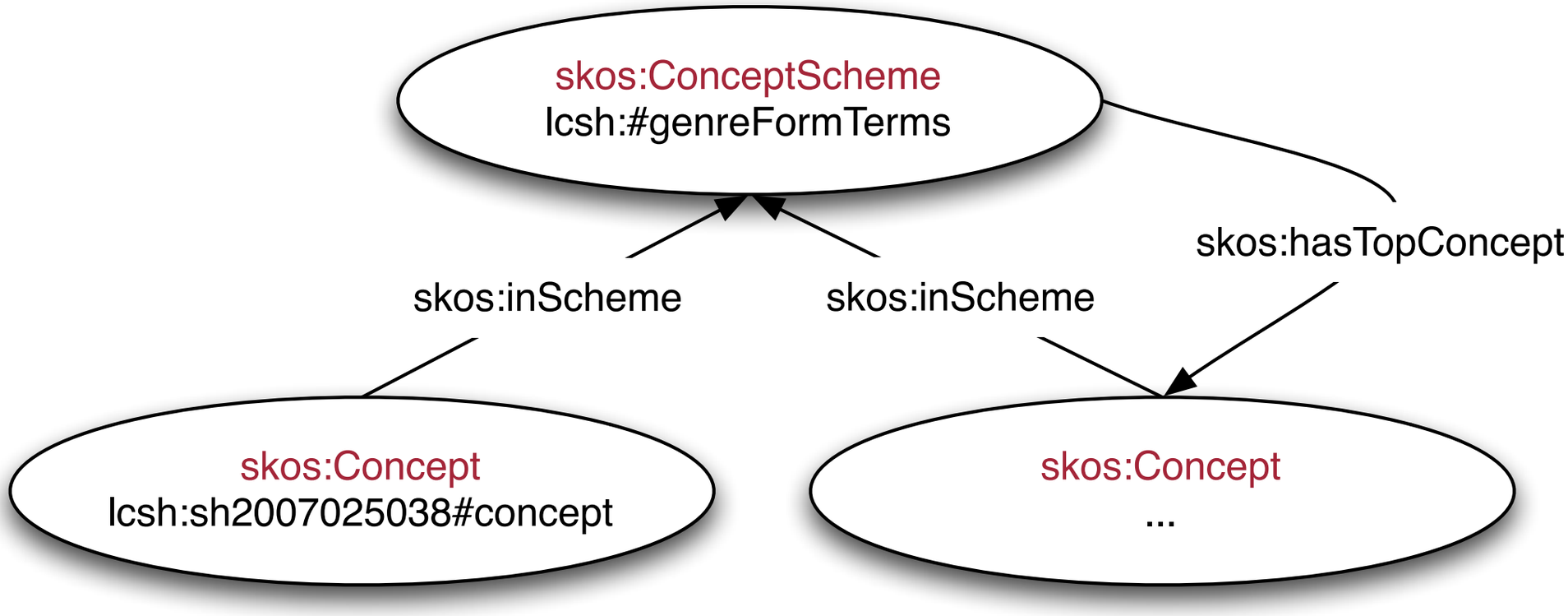
SKOS Documentary Notes

- Add further human-readable documentation
  - skos:scopeNote: info about intended meaning
  - skos:definition: complete explanation of meaning
  - skos:example: example concept use



skos:ConceptScheme

- Allow the organization of skos:Concepts in some Knowledge Organization Scheme (KOS)



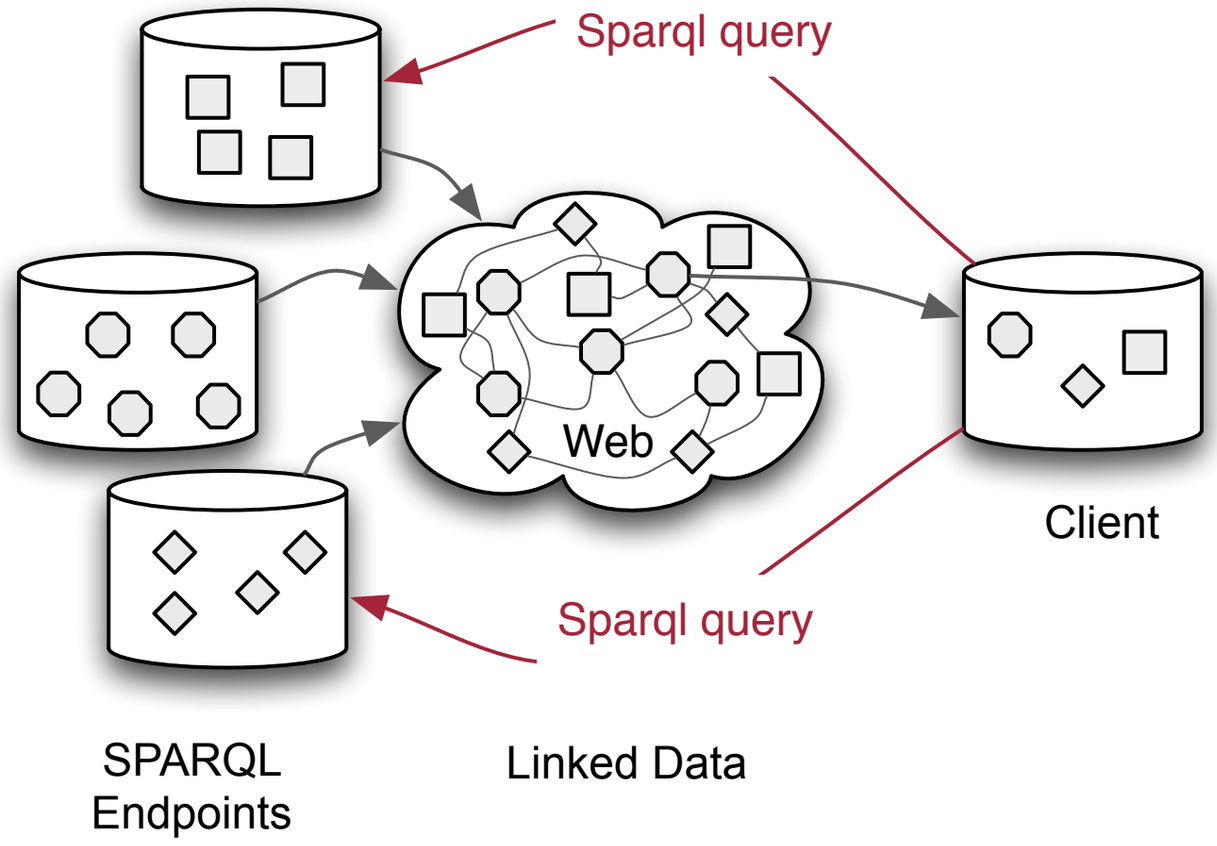
A Real-World Example

- NY Times topics: http://topics.nytimes.com/
- Dereference and analyze "Jack Nicholson" @ New York Times
  - http://data.nytimes.com
  - http://data.nytimes.com/N57614112774312665 13

SPARQL

What is SPARQL?

- SPARQL is a query language for accessing RDF data
- SPARQL is a protocol that defines how queries and results can be transported over a network



Example Dataset

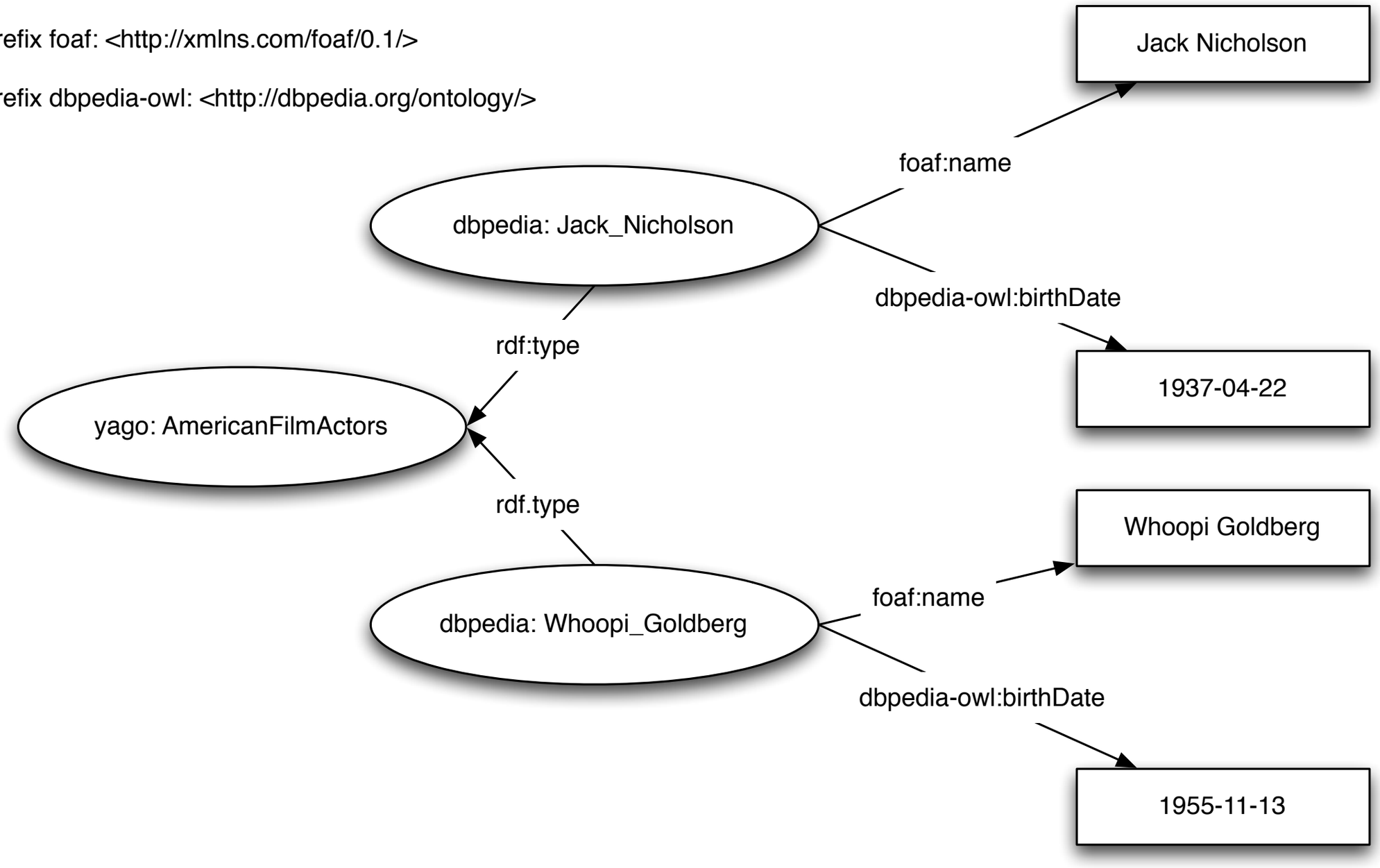
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

@prefix dbpedia: <http://dbpedia.org/resource/>

@prefix yago: <http://dbpedia.org/class/yago/>

@prefix foaf: <http://xmlns.com/foaf/0.1/>

@prefix dbpedia-owl: <http://dbpedia.org/ontology/>



A Simple SPARQL Query

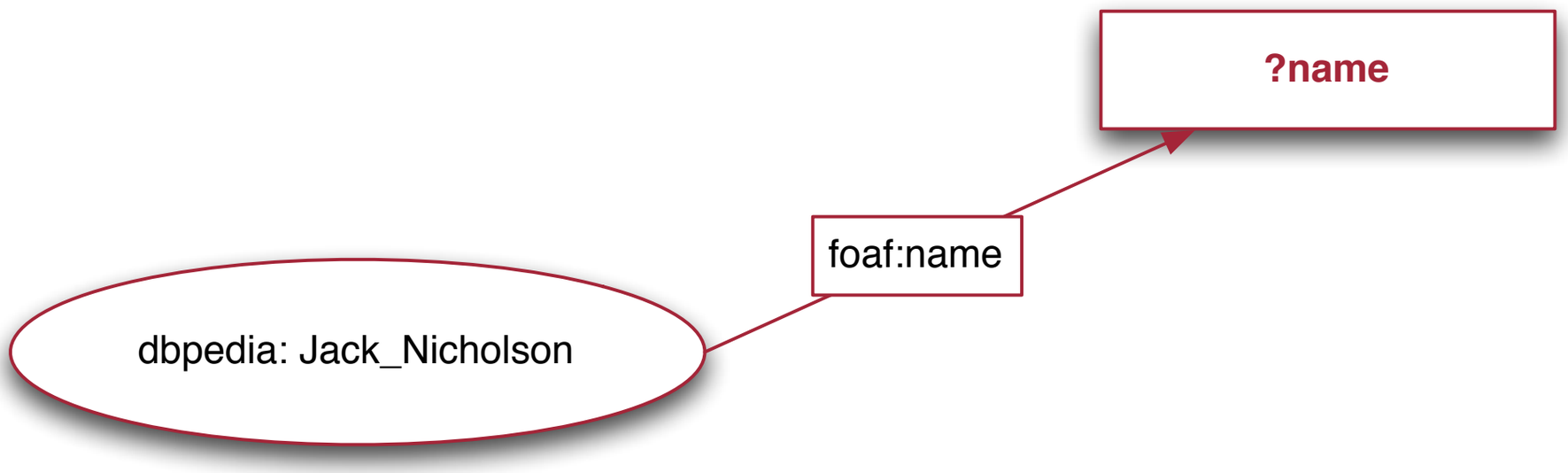
```
SELECT ?name
WHERE {
    dbpedia:Jack_Nicholson foaf:name ?name
}
```

name
“Jack Nicholson”

Simple Query Illustrated

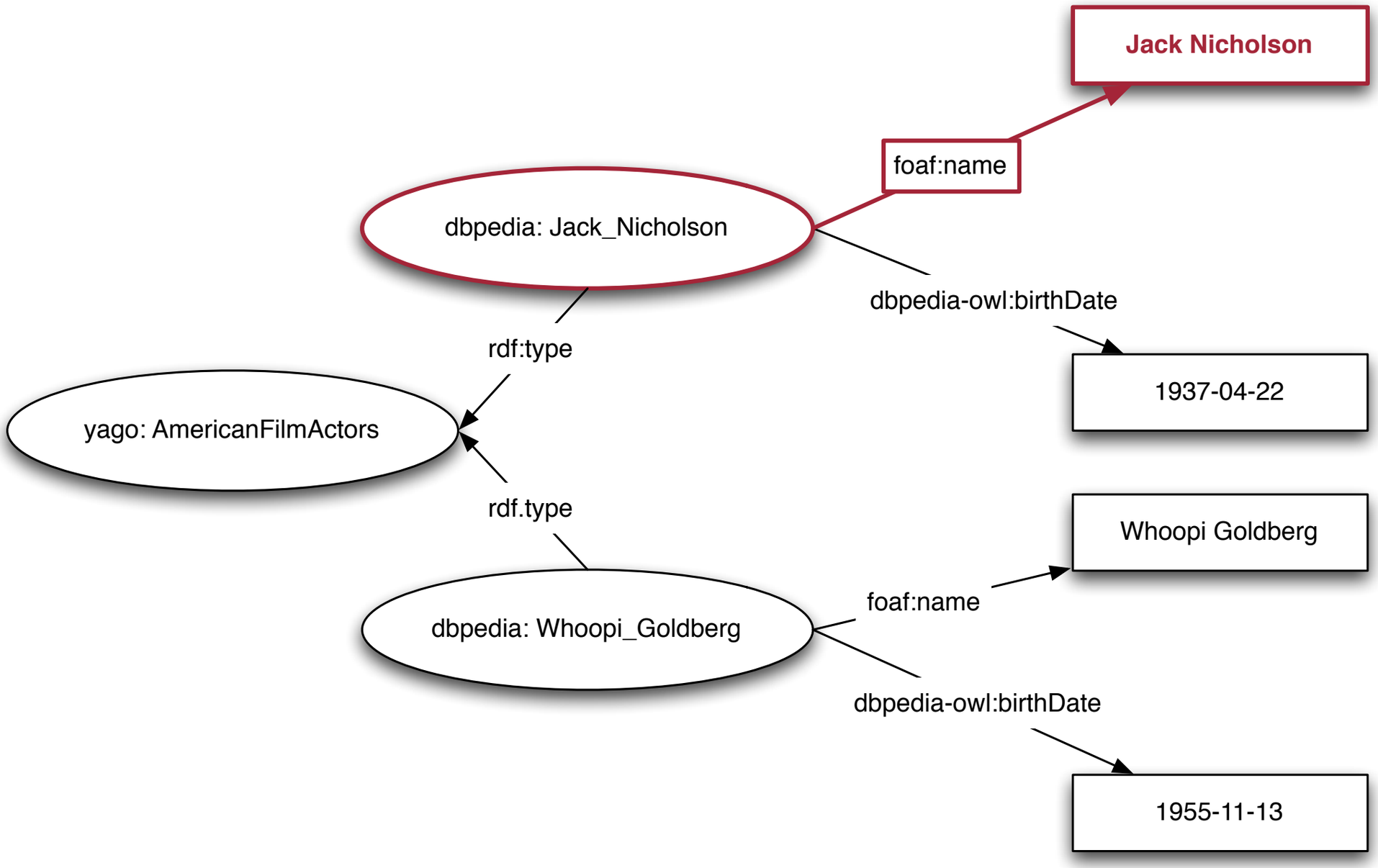
- In SPARQL we formulate triple/graph patterns

```
SELECT ?name
WHERE {
      dbpedia:Jack Nicholson foaf:name ?name
}
```



Simple Query Illustrated

- Patterns are matched against the dataset



Querying Multiple Values

```
SELECT ?name ?birthdate
WHERE {
      ?x rdf:type yago:AmericanFilmActors .
      ?x foaf:name ?name .
      ?x dbpedia-owl:birthDate ?birthdate .
}
```

name	birthdate
“Jack Nicholson”	1937-04-22
“Whoopi Goldberg”	1955-11-13

```
SELECT ?name ?birthdate
WHERE {
     ?x rdf:type yago:AmericanFilmActors .
     ?x foaf:name ?name .
     ?x dbpedia-owl:birthDate ?birthdate .
}
```

Exercise: draw the graph pattern

Prefixes in SPARQL queries

- The DBPedia SPARQL Web interface knows about prefix/namespace mappings
- But in general SPARQL endpoints don't

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX dbpedia: <http://dbpedia.org/resource/>

SELECT ?name
WHERE {
    dbpedia:Jack_Nicholson foaf:name ?name
}
```

SPARQL Filters

- ...restrict the solutions of a graph pattern match according to a given expression
- ...eliminate solutions that, when substituted into the expression, result in boolean false

```
SELECT ?name
WHERE {
     ?x foaf:name ?name .
     FILTER regex(?name, “Nicholson”)
}
```

name
“Jack Nicholson”

SPARQL Filters Example

```
SELECT ?name ?birthdate
WHERE {
     ?x rdf:type yago:AmericanFilmActors .
     ?x foaf:name ?name .
     ?x dbpedia-owl:birthDate ?birthdate .
     FILTER(?birthdate > "1950-01-01T00:00:00Z"^^xsd:dateTime)
}
```

name	birthdate
“Whoopi Goldberg”	1955-11-13

More SPARQL Filters

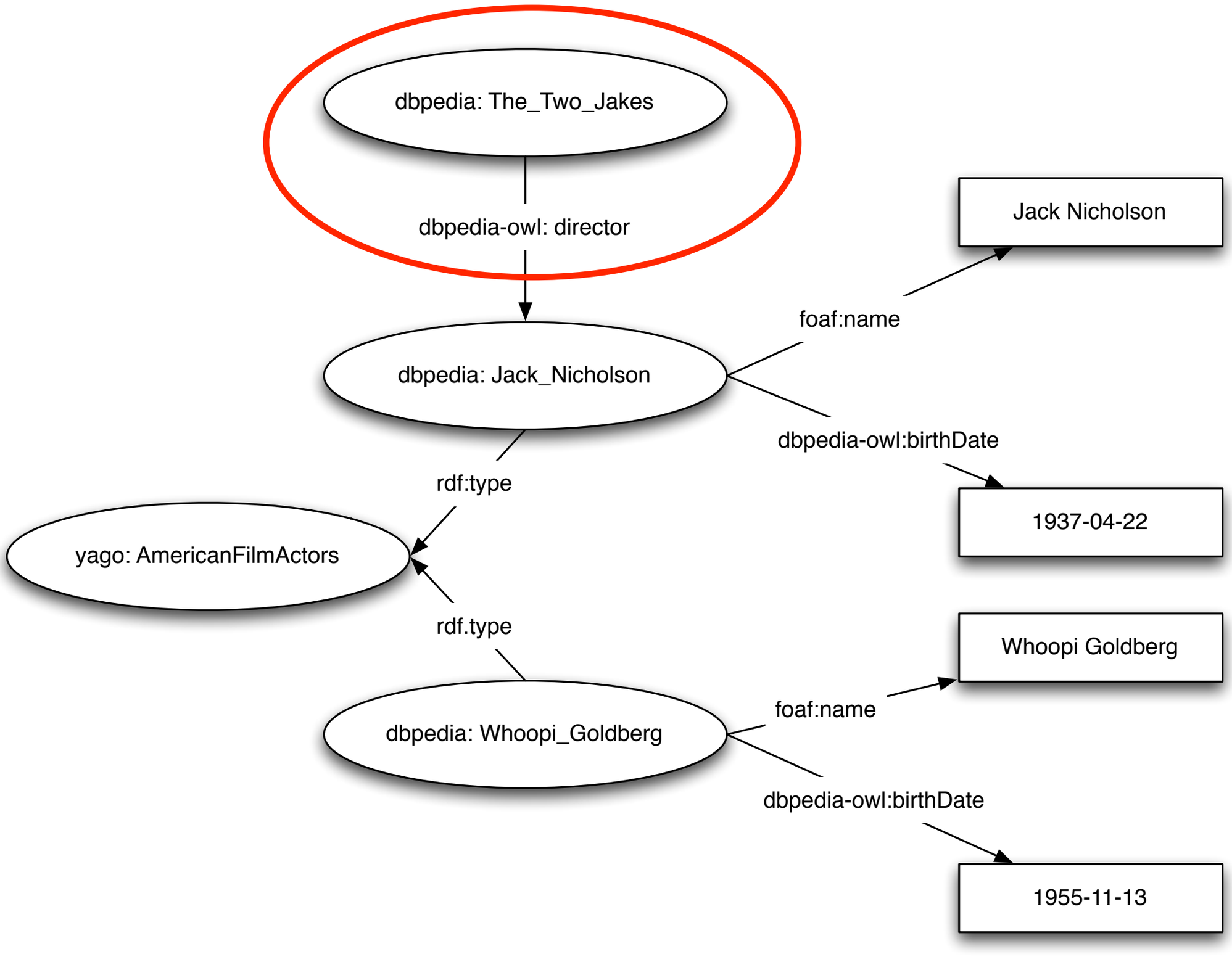
- SPARQL offers lots of FILTERing possibilities
  - Boolean operators ! || &&
  - Comparison operators = != > < >= <=
  - Arithmetic operators * / - +
  - RDF element operators bound(), isURI(), LANG(), STR()

further info: http://www.w3.org/TR/rdf-sparql-query/#tests

Optional Graph Patterns

- In a standard SPARQL query the entire graph pattern must match in order to retrieve a result
- But we cannot always assume complete structures
  - not all actors have birthdays
  - not all movies have actors
- SPARQL allows to formulate queries that include information in the solution if it is available, but does not reject the solution if parts are missing

Optional Graph Patterns



Optional Graph Patterns

```
SELECT ?name ?directed_movie
WHERE {
      ?x rdf:type yago:AmericanFilmActors .
      ?x foaf:name ?name .
      ?directed_movie dbpedia-owl:director ?x .
}
```

name	directed_movie
"Jack Nicholson"	dbpedia:The_Two_Jakes

Exercise: we lost Whoopi. Why?

Optional Graph Patterns

```
SELECT ?name ?directed_movie
WHERE {
      ?x rdf:type yago:AmericanFilmActors .
      ?x foaf:name ?name .
      OPTIONAL {
            ?directed movie dbpedia-owl:director ?x .
      }
}
```

name	directed_movie
“Jack Nicholson”	dbpedia:The_Two_Jakes
“Whoopi Goldberg”	

SPARQL UNION

- Sometimes you want to express „or“ in a graph pattern
- „The graph should match this OR that pattern“
- With the UNION keyword we can define alternative matching graph patterns

SPARQL UNION Example

```
SELECT ?name ?directed_movie
WHERE {

    {?x rdf:type yago:AmericanFilmActors }
    UNION
    {?x rdf:type yago:GermanFilmActors }

    ?x foaf:name ?name .

    OPTIONAL {
        ?directed_movie dbpedia-owl:director ?x .
    }
}
```

SPARQL Solution Modifiers

- The results returned by a query are by default unordered
- SPARQL defines the following solution modifiers
  - ORDER BY - reorder the solution sequence
  - DISTINCT - avoid duplicate solutions
  - OFFSET - start after a certain number of solutions
  - LIMIT - limit the output to a number of solutions

SPARQL Solution Modifiers

```
SELECT ?name
WHERE {
     ?x a foaf:Person .
     ?x foaf:name ?name .
}
LIMIT 100
```

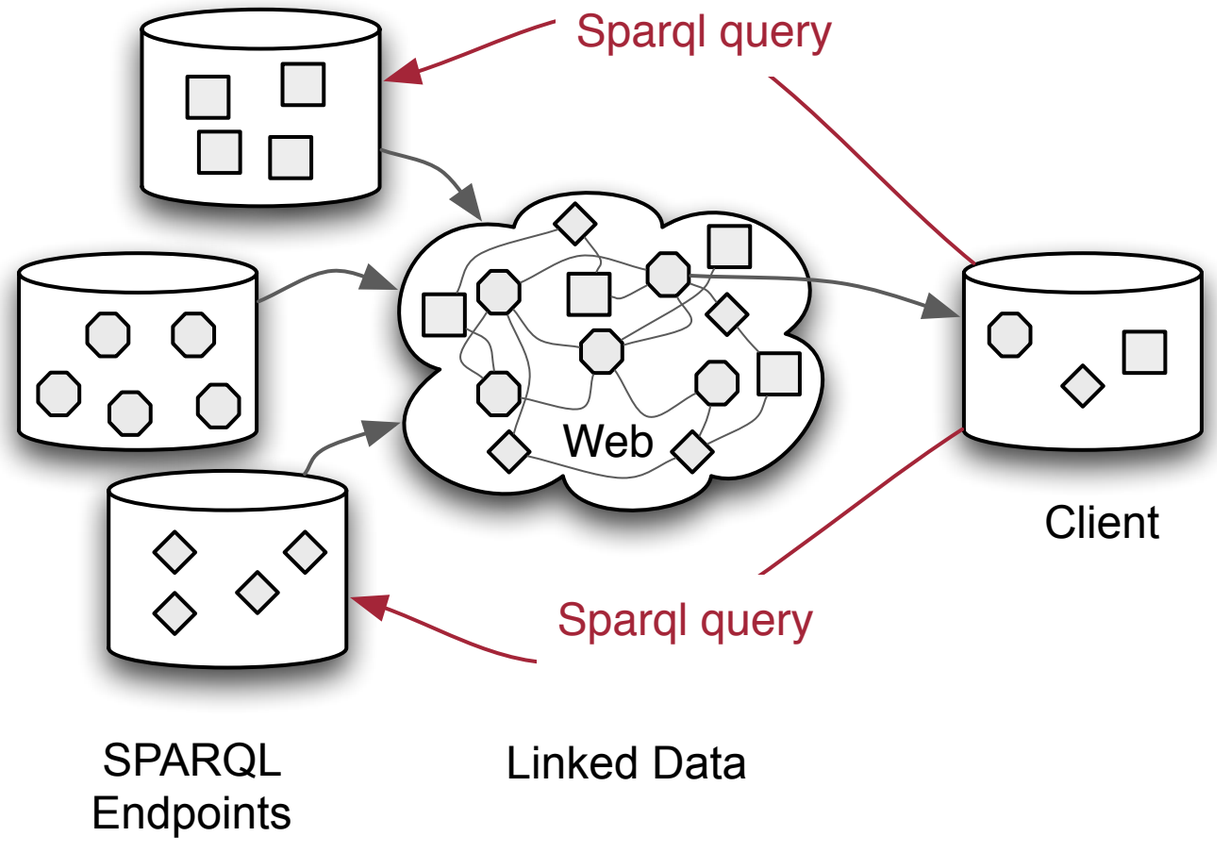
```
SELECT ?person (count(DISTINCT ?spouse) as ?spouses)
where {
     ?person a yago:AmericanFilmActors .
     ?person dbpprop:spouse ?spouse .
}
ORDER BY DESC(?spouses)LIMIT 100
```

Further SPARQL features

- We only covered SELECT
- There are also the CONSTRUCT, ASK, and DESCRIBE query forms
- The other forms are useful in practice!
- Please read the SPARQL Primer: http://www.w3.org/TR/rdf-sparql-query/

What is SPARQL?

- SPARQL is a query language for accessing RDF data
- SPARQL is a protocol that defines how queries and results can be transported over a network



SPARQL Protocol for RDF

- SPARQL also defines how queries and and results can be transported over a network
- Bindings for HTTP
  http://www.w3.org/TR/rdf-sparql-protocol/

SPARQL Protocol - HTTP

- Queries are sent to an endpoint using
  - HTTP GET (default)
  - HTTP POST (if encoded query string exceed limits)
- Results are returned either as
  - SPARQL Results Document (SELECT and ASK)
  - Serialized RDF Graph (CONSTRUCT and DESCRIBE)

SPARQL Protocol - Example

```
SELECT ?name
WHERE {
    dbpedia:Jack_Nicholson foaf:name ?name
}
```

URI-encoded query string (`EncodedQuery`)

```
SELECT%20?name%0AWHERE%20%7B%0A
%20%20%20%20dbpedia:Jack_Nicholson
%20foaf:name%20?name%0A%7D
```

SPARQL Protocol - Example

```
GET /sparql/?query=EncodedQuery HTTP/1.1
Host: dbpedia.org
```

Using CURL:

```
curl -v http://dbpedia.org/sparql/?
query=SELECT%20?name%0AWHERE%20%7B%0A
%20%20%20%20dbpedia:Jack_Nicholson
%20foaf:name%20?name%0A%7D
```

SPARQL Protocol - Example

```
<sparql xmlns="...”>
<head>
    <variable name="name"/>
</head>
<results distinct="false" ordered="true">
    <result>
        <binding name="name">
            <literal xml:lang="en">
            Jack Nicholson
            </literal>
        </binding>
    </result>
</results>
```

QUESTIONS?