(Technical) Introduction to Linked Data

ERIK VAN DEN BERGH

FAIR DATA SPRINT 4TH FEBRUARY 2020

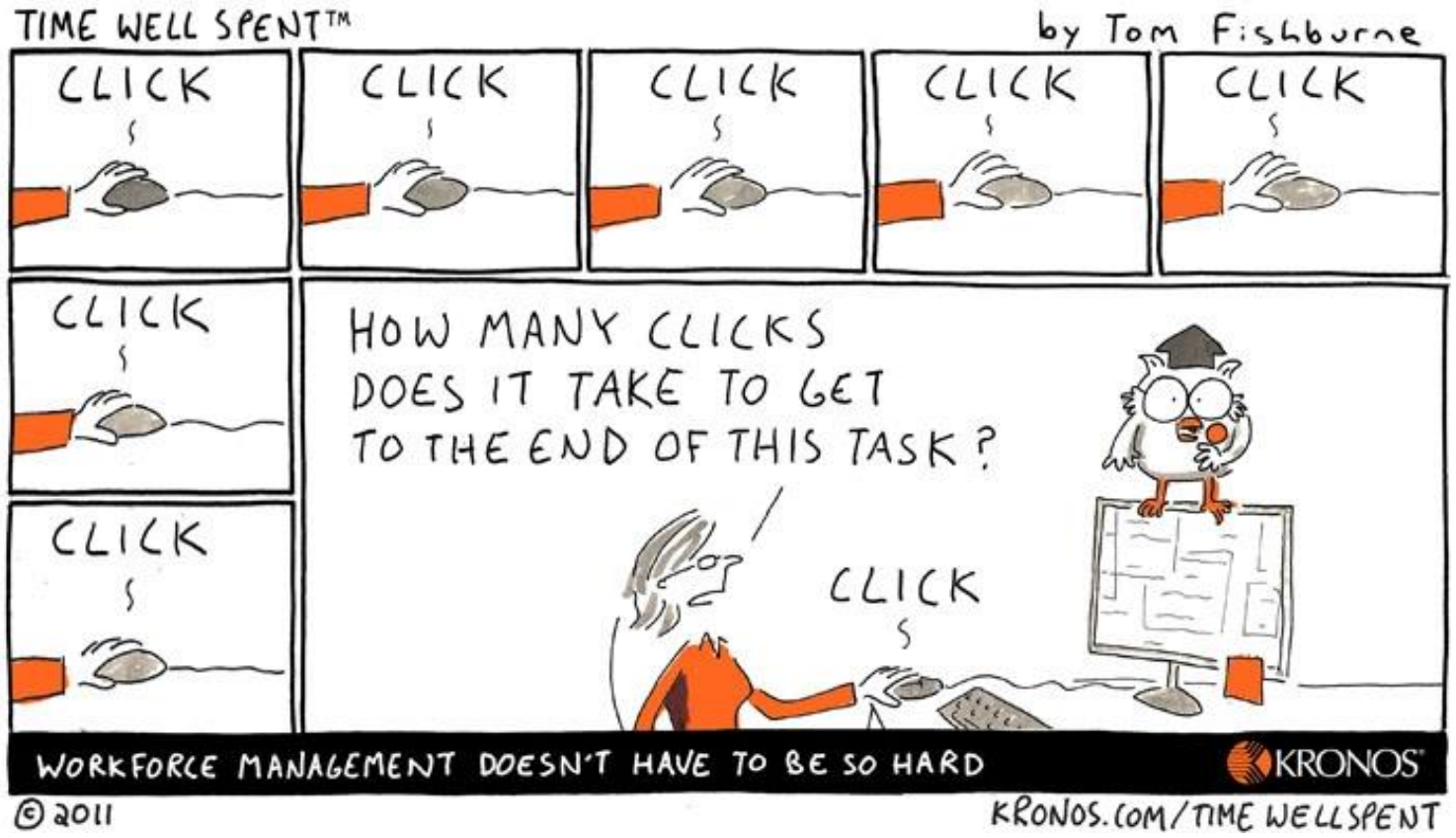
Currently used (Life Science) web resources

Link

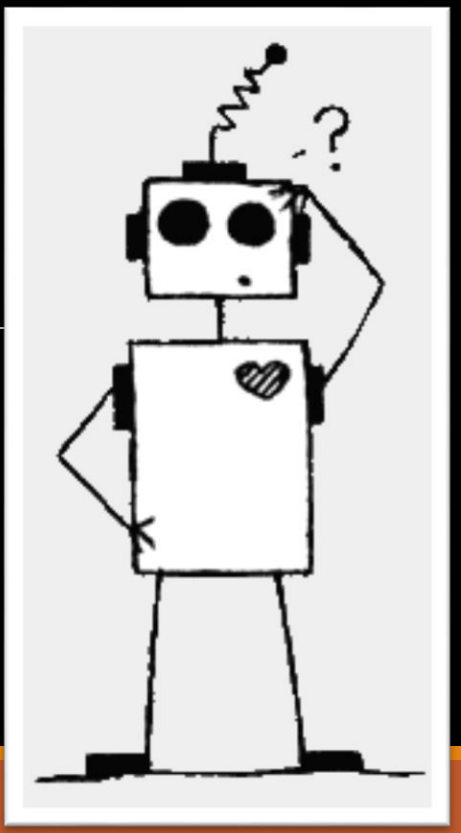
Link

Resource through the “old” internet

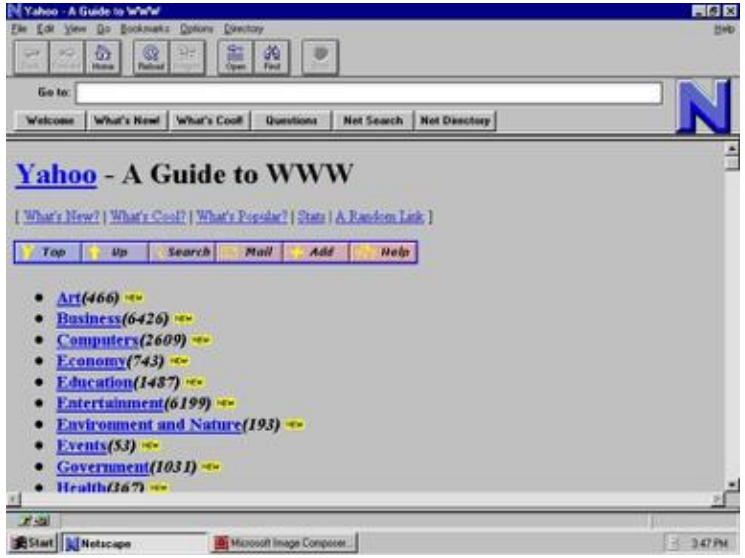
100 clicks later….



Machine readable?



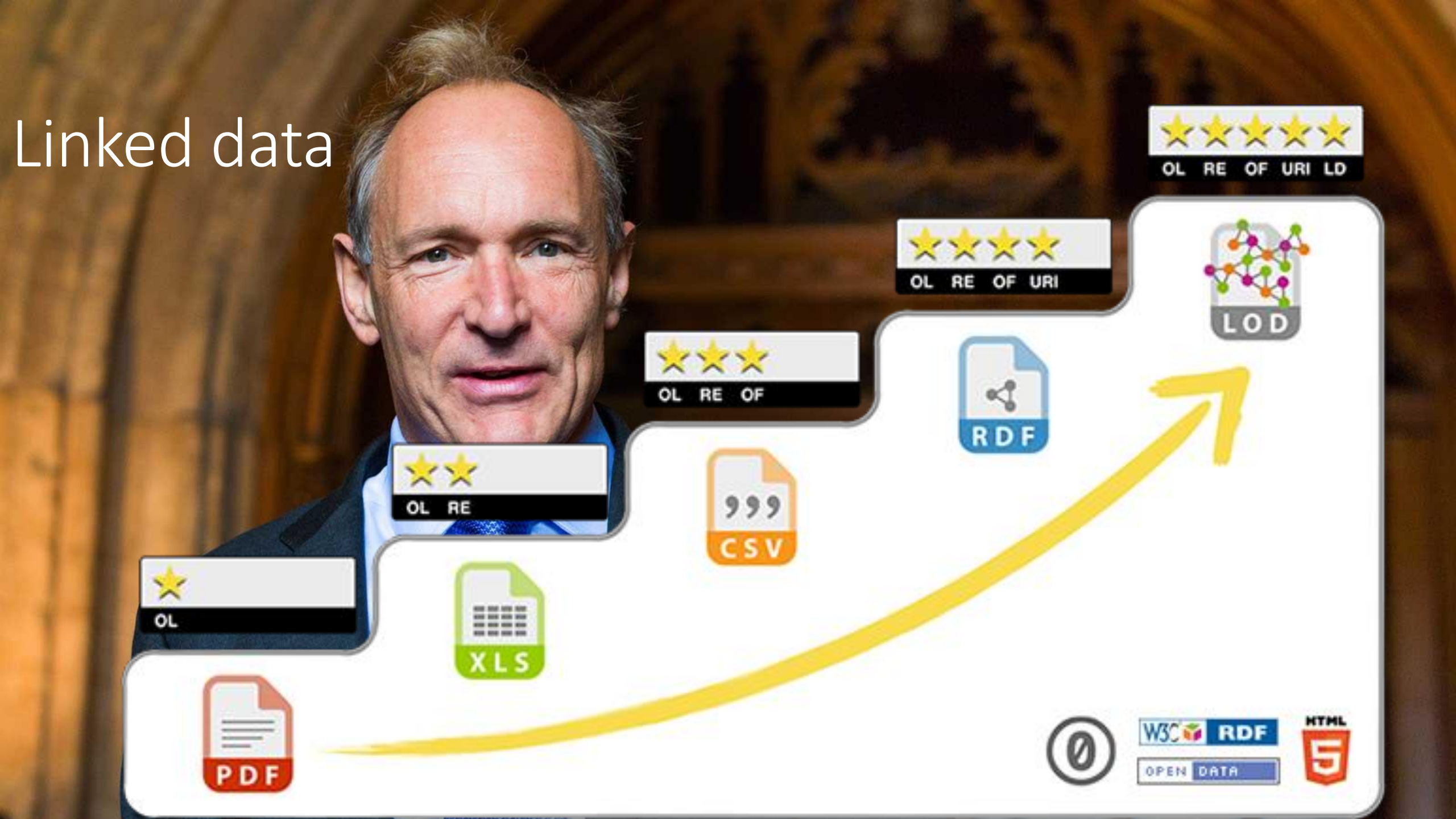
WEB 1.0



WEB 3.0

WEB 2.0

Linked data


OL
OL RE
OL RE OF
OL RE OF URI
OL RE OF URI LD
PDF
XLS
CSV
RDF
LOD
W3C RDF
OPEN DATA
HTML 5

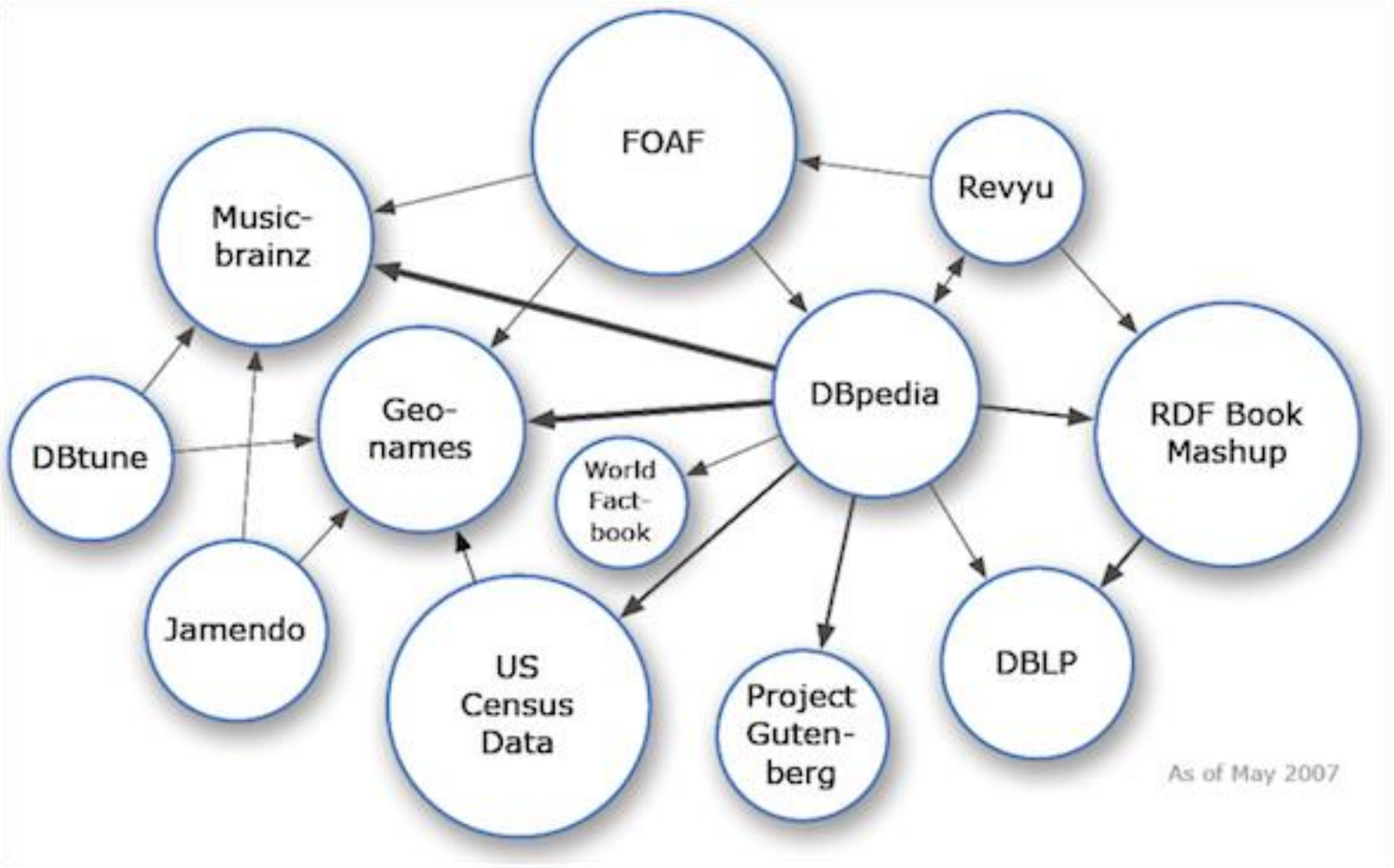
Internet of things for the general public

Where does the research come in?

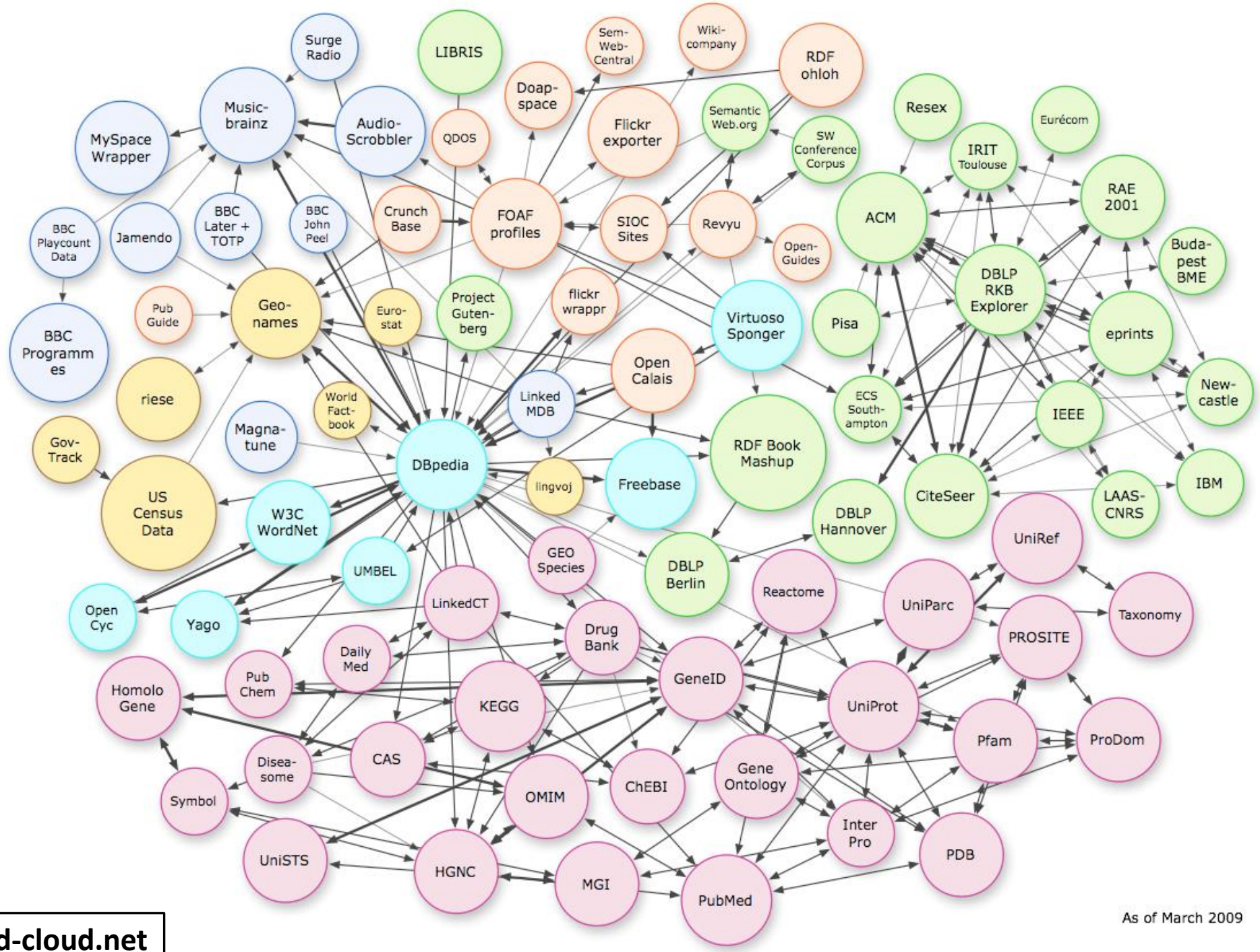
The databases behind them…

Devices can access the databases but so can we!

What is out there?

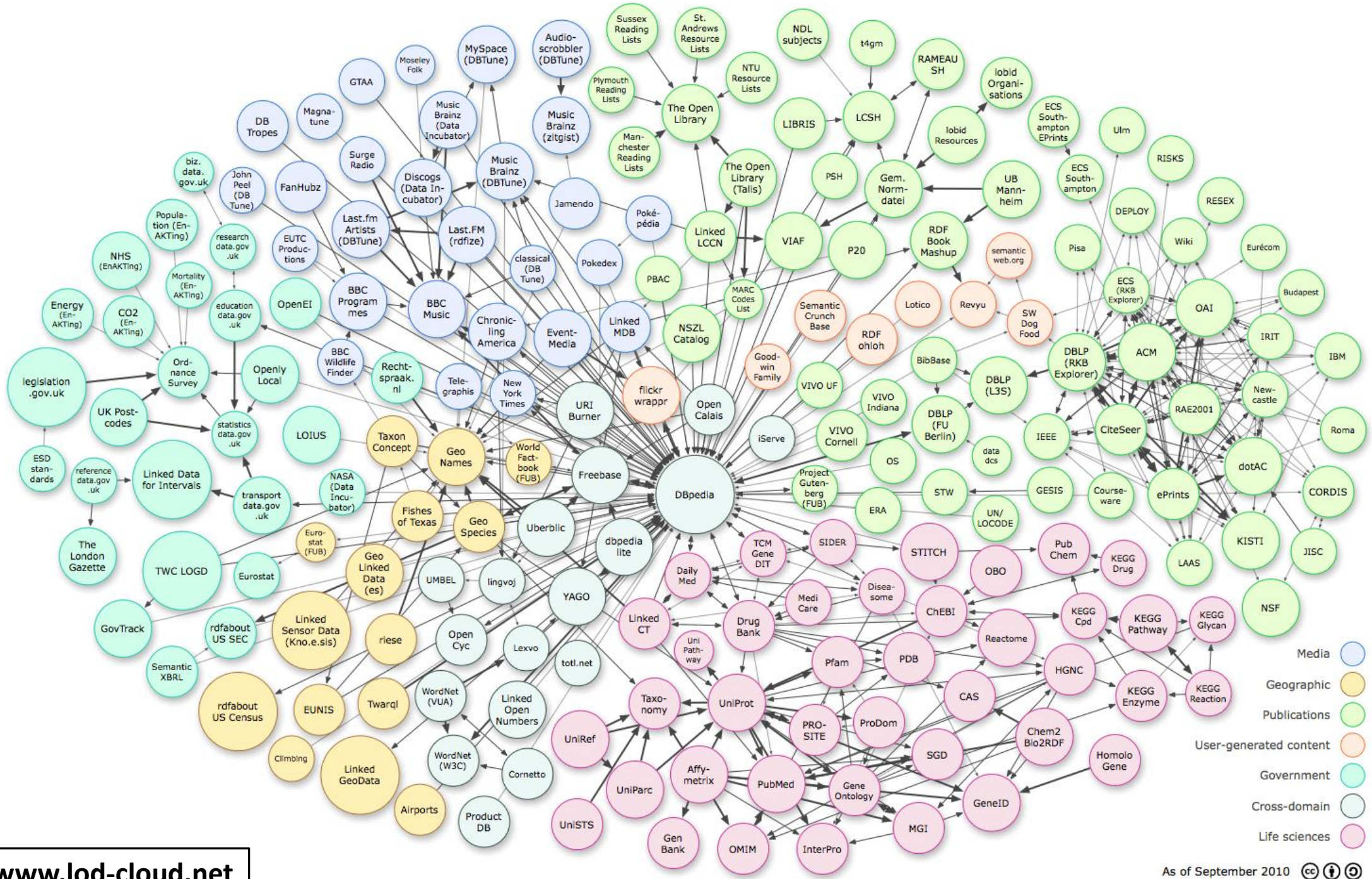


https://www.lod-cloud.net

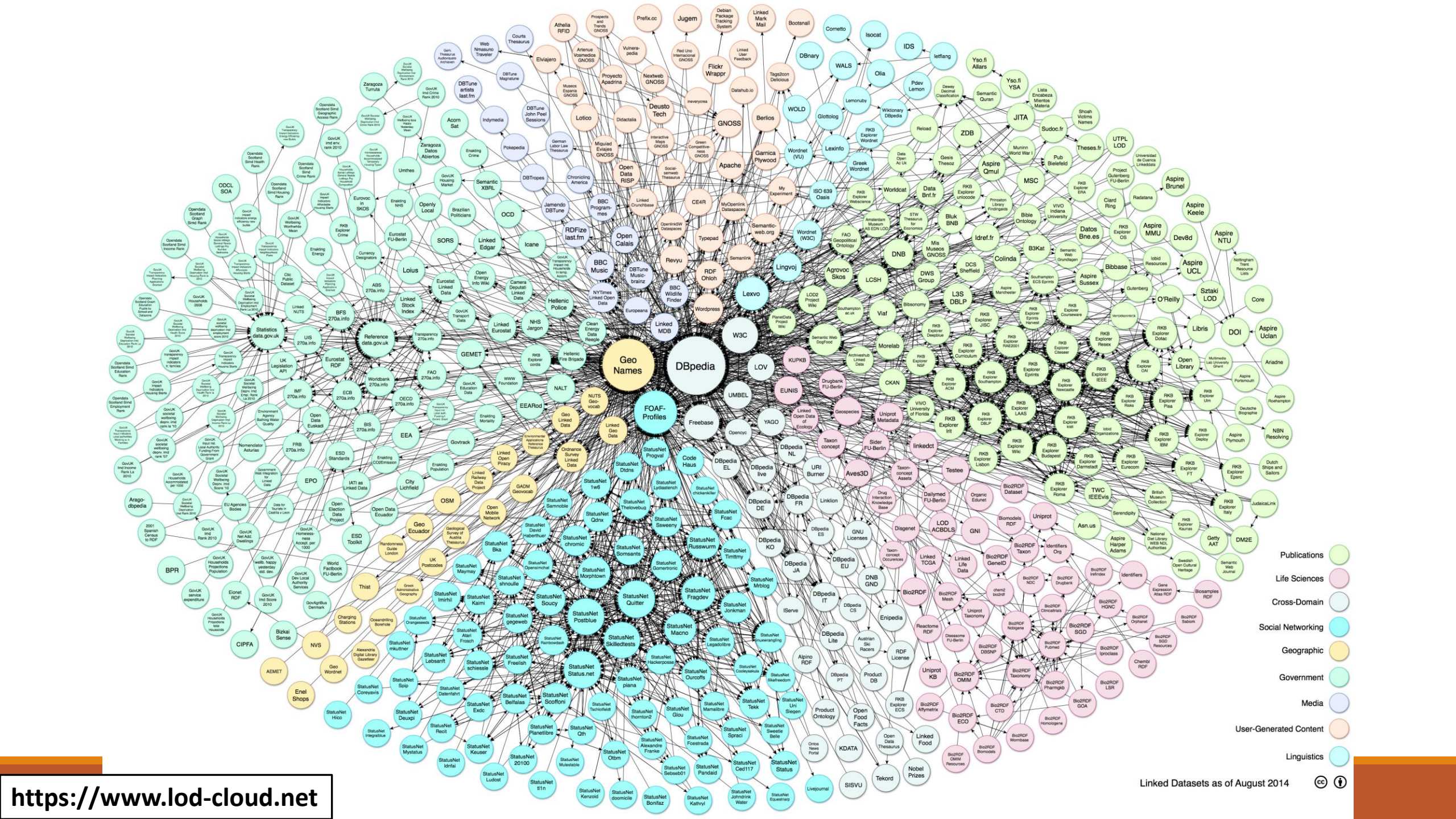


As of March 2009

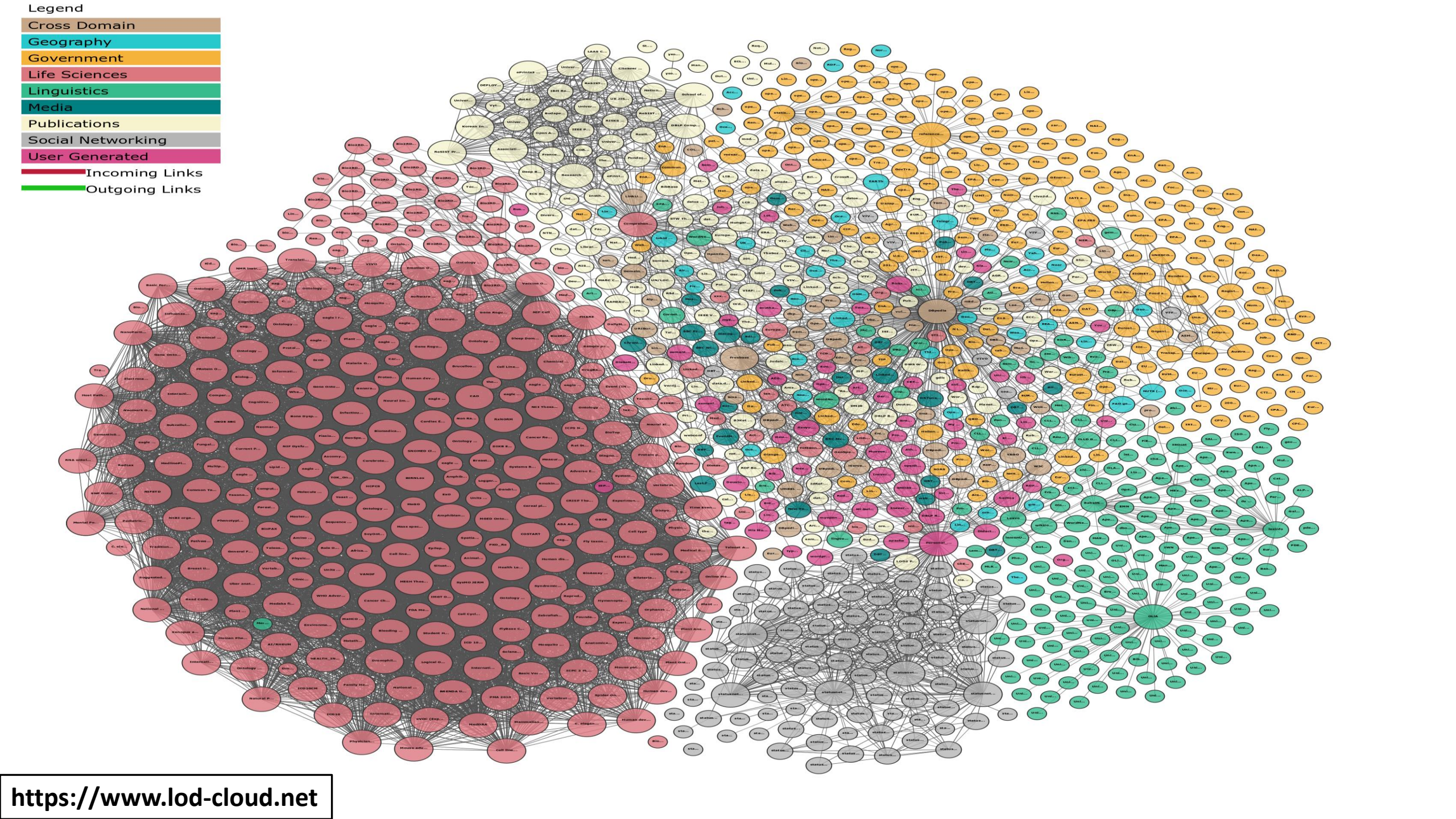
https://www.lod-cloud.net



https://www.lod-cloud.net



https://www.lod-cloud.net



https://www.lod-cloud.net

How to access these datasets?

The current standard for sharing data is through linked data represented as RDF / Triples

Resource Description Framework (RDF)

A framework for describing resources

Resources: Everything that can be represented with URLs

Description: Statements about the properties of such resources

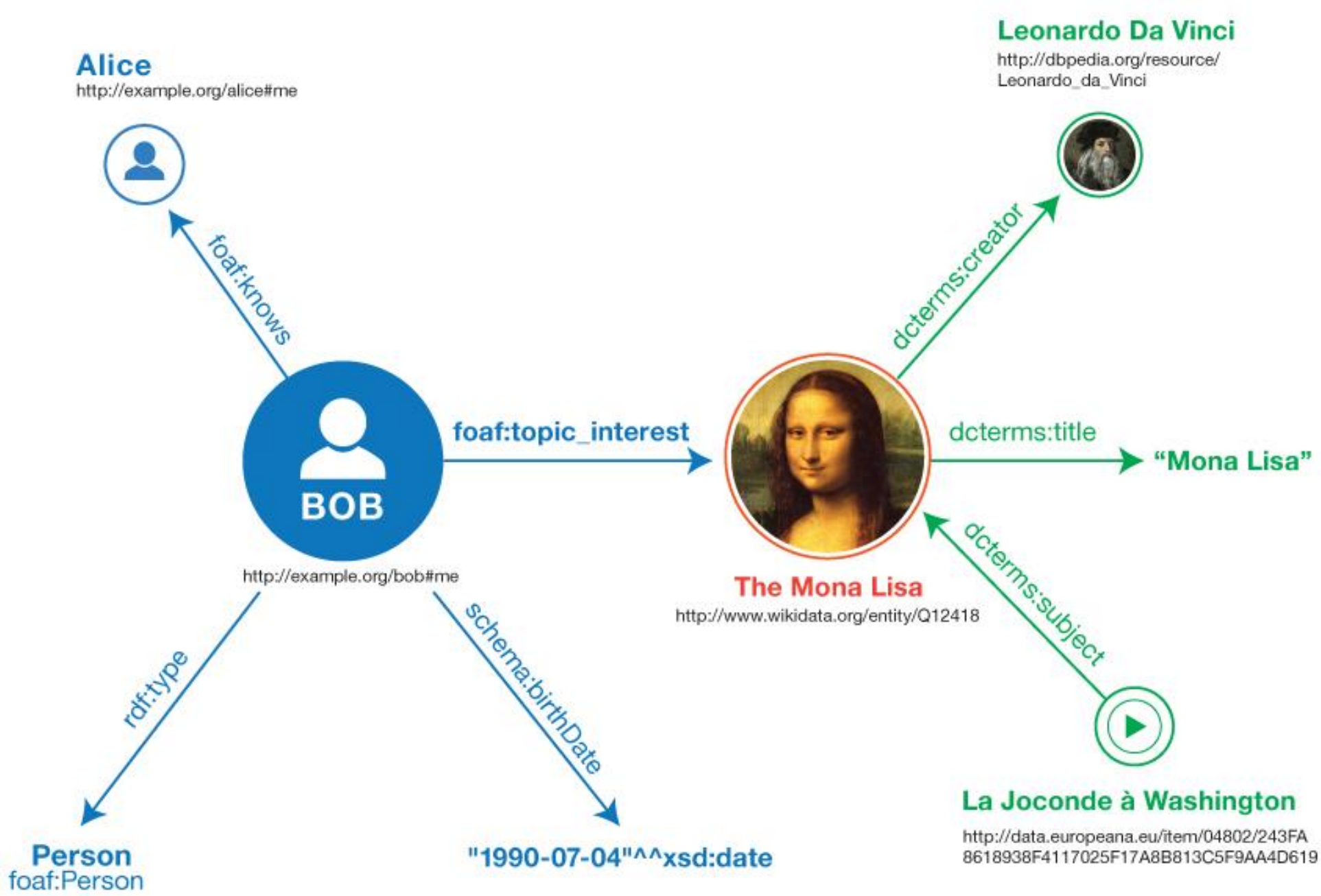
Framework: A common model describing the data

Based on URLs (http://www….) using triples

Triples


SUBJECT
PREDICATE
OBJECT
I
KNOWS
John
Subject
IRI
Predicate
IRI
Object
IRI/Value

Triples


Alice
http://example.org/alice#me
Leonardo Da Vinci
http://dbpedia.org/resource/Leonardo_da_Vinci
foaf:knows
dcterms:creator
BOB
http://example.org/bob#me
foaf:topic_interest
dcterms:title
"Mona Lisa"
The Mona Lisa
http://www.wikidata.org/entity/Q12418
dcterms:subject
rdf:type
schema:birthDate
Person
foaf:Person
"1990-07-04"^^xsd:date
La Joconde à Washington
http://data.europeana.eu/item/04802/243FA8618938F4117025F17A8B813C5F9AA4D619

Speaking the same language

Ontology:

- is a term in philosophy: ``theory of existence''.
- is an explicit specification of conceptualization.
- is a body of knowledge describing a knowledge domain.

Conoce, kent, has been introduced to, met, talked to, connaître

All means the same but are different for a computer

foaf:knows is a defined standard for this relation

Knowing what it means…

Ontology URLs are often resolvable (meaning they point to a real web address)

foaf: > short cut for > http://xmlns.com/foaf/0.1/

foaf:Person >> http://xmlns.com/foaf/0.1/Person

Class: foaf:Person

Person - A person.

Status:	stable
Properties include:	plan surname geekcode pastProject lastName family_name publications currentProject familyName firstName workInfoHomepage myersBriggs schoolHomepage img workplaceHomepage knows
Used with:	knows
Subclass Of	Agent Spatial Thing
Disjoint With:	Project Organization

The `Person` class represents people. Something is a `Person` if it is a person. We don't nitpic about whether they're alive, dead, real, or imaginary. The `Person` class is a sub-class of the `Agent` class, since all people are considered 'agents' in FOAF.

FOAF

http://xmlns.com/foaf/0.1/knows

http://example.com/person/jasper

http://example.com/person/John

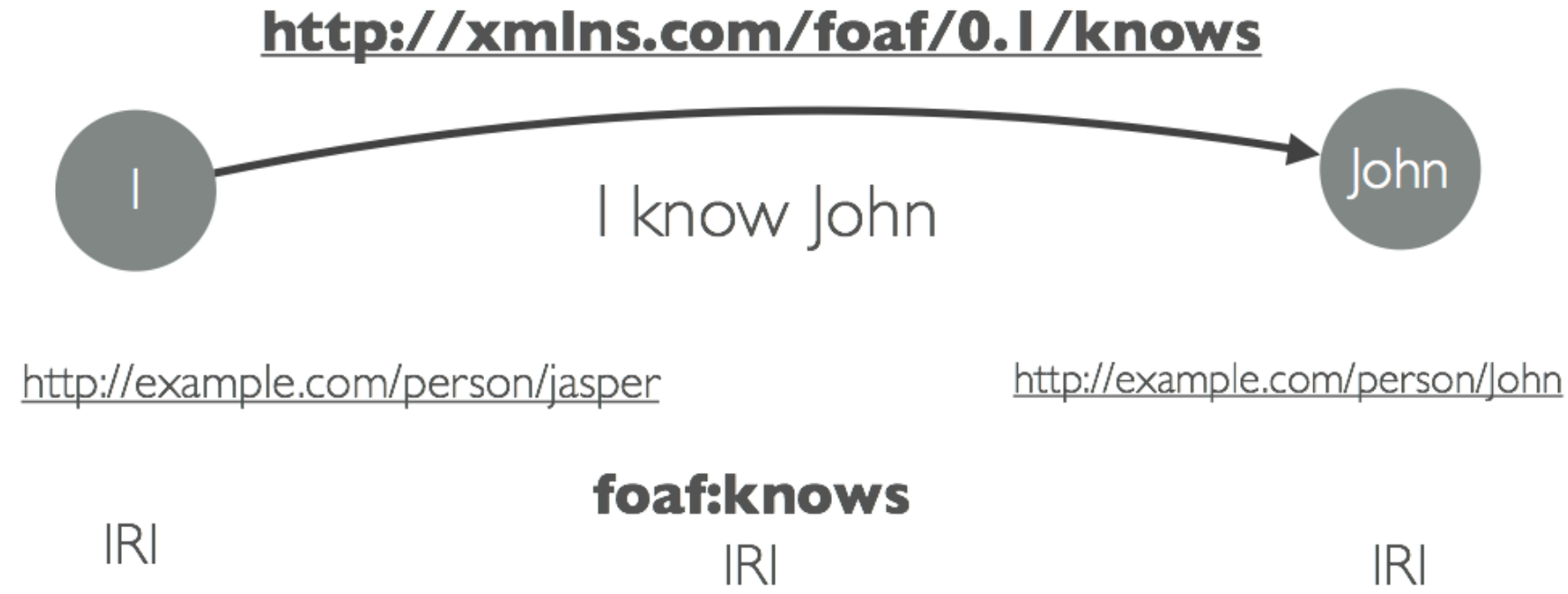
foaf:knows

IRI IRI IRI

http://dbpedia.org/ontology/hometown

IRI IRI Value

Writing triples



<http://example.com/person/jasper> <http://xmlns.com/foaf/0.1/knows> <http://example.com/person/John>

Writing triples

<http://example.com/person/jasper> <http://xmlns.com/foaf/0.1/knows> <http://example.com/person/John> .
<http://example.com/person/jasper> <http://dbpedia.org/ontology/hometown> “wageningen” .

Inventing your own

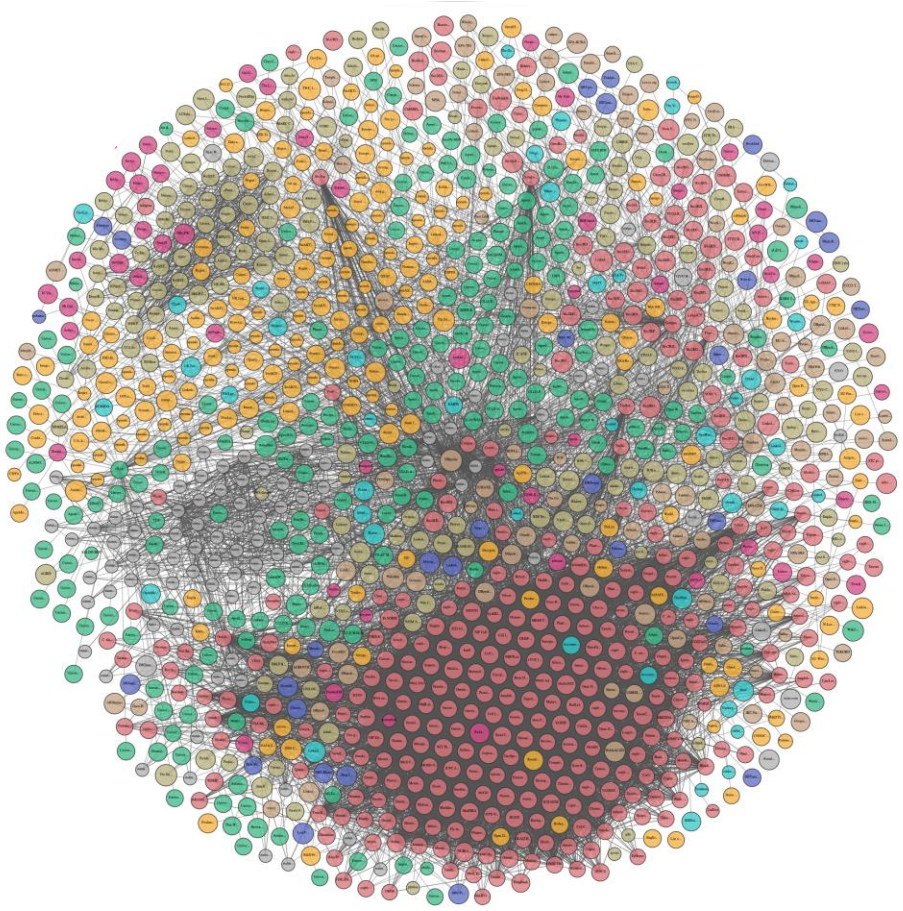
<http://example.com/person/jasper> <http://xmlns.com/foaf/0.1/knows> <http://example.com/person/John> .
<http://example.com/person/jasper> <http://dbpedia.org/ontology/hometown> “wageningen” .

<http://example.com/person/jasper> <http://xmlns.com/foaf/0.1/familyName> “Koehorst” .

<http://example.com/person/jasper> **<http://myontology.com/pet>** <http://example.com/animal/puk> .

The cloud

- Each node is an RDF endpoint (public / private)
- You can query from one database to another
  - **Federated queries**
- You can start from your own private database and access a public database in one query

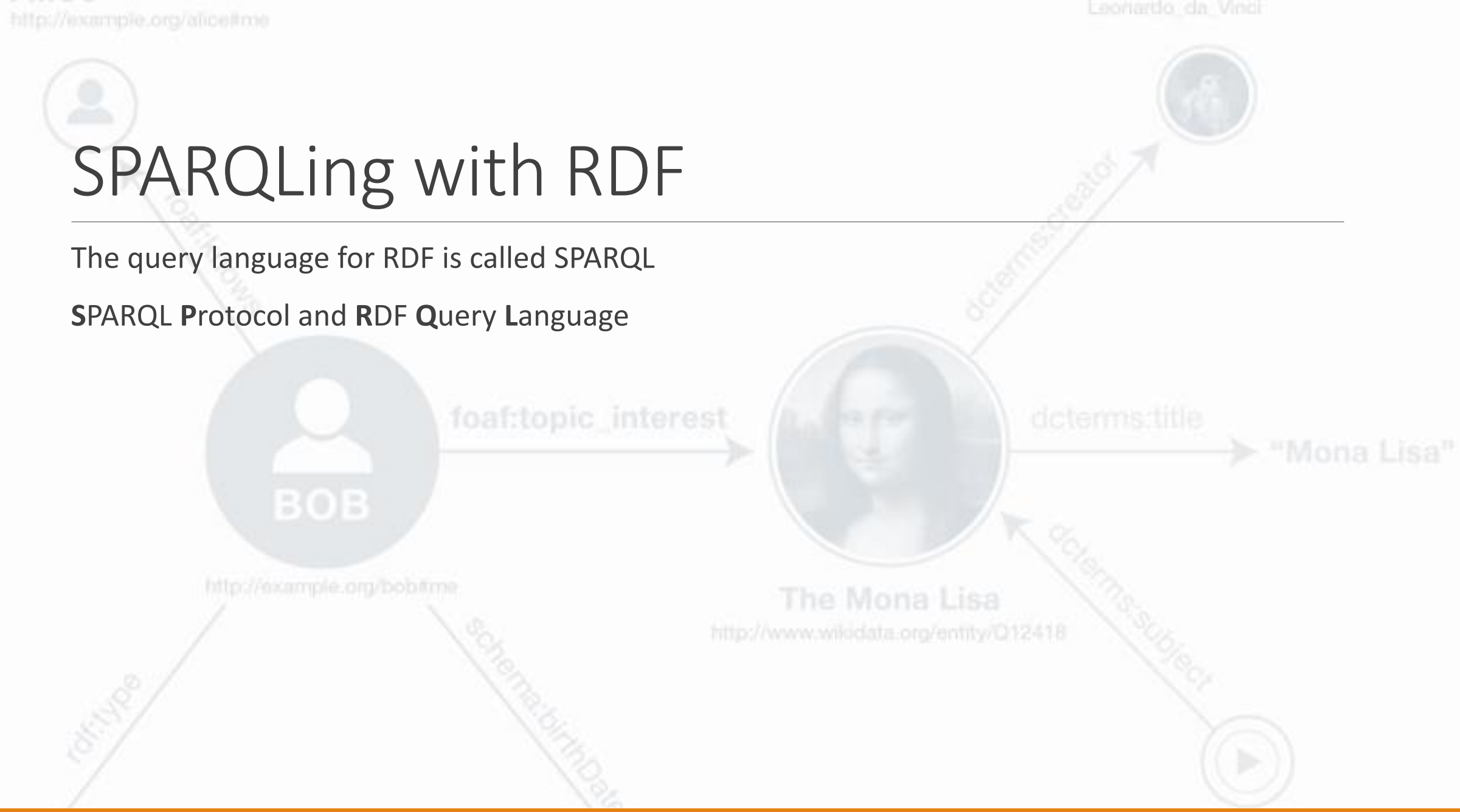


- Query? >>> **SPARQL**!

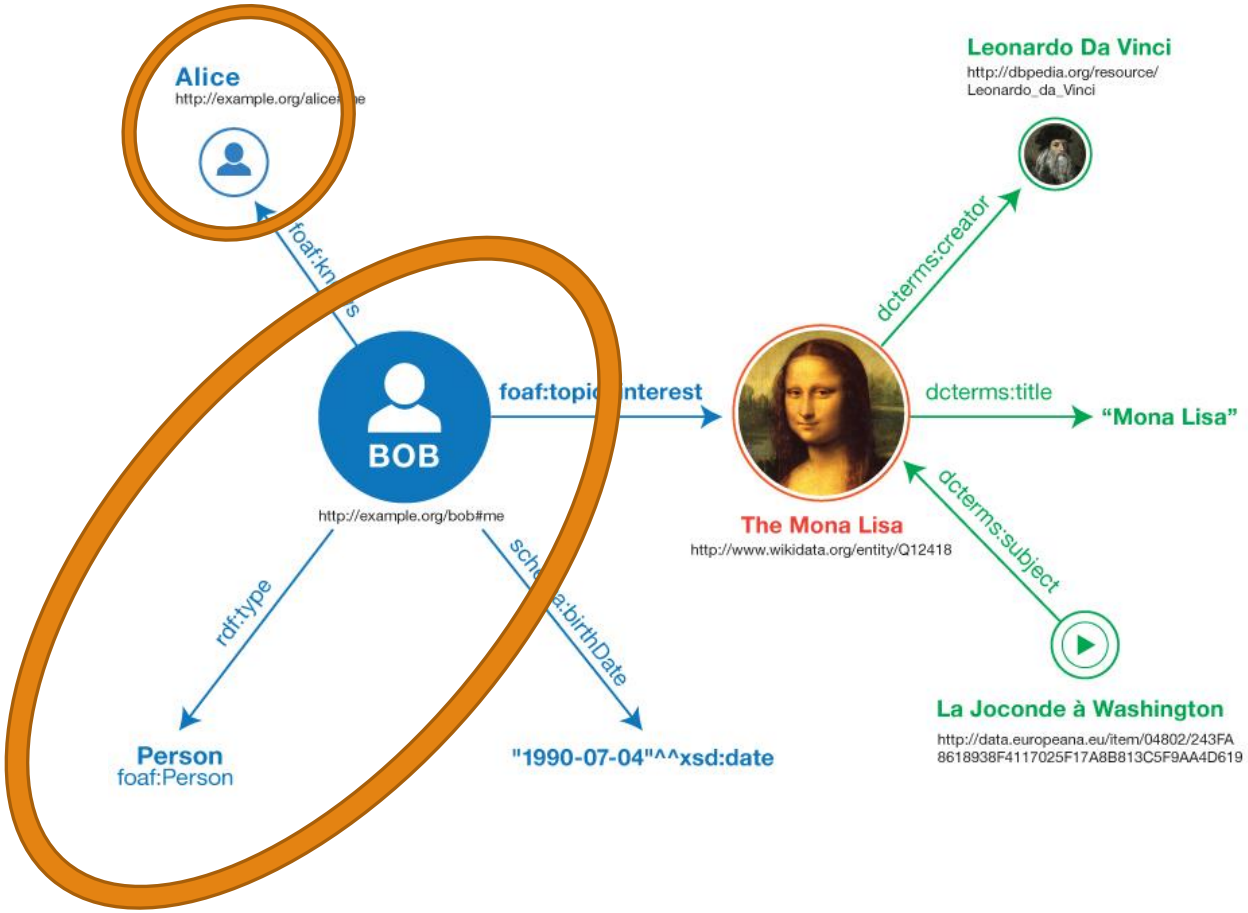
SPARQLing with RDF

The query language for RDF is called SPARQL

SPARQL **P**rotocol and **R**DF **Q**uery **L**anguage



Who is human?




```
PREFIX foaf:<http://xmlns.com/foaf/0.1/>
SELECT ?person
WHERE {
     ?person a foaf:Person .
}
```

?person

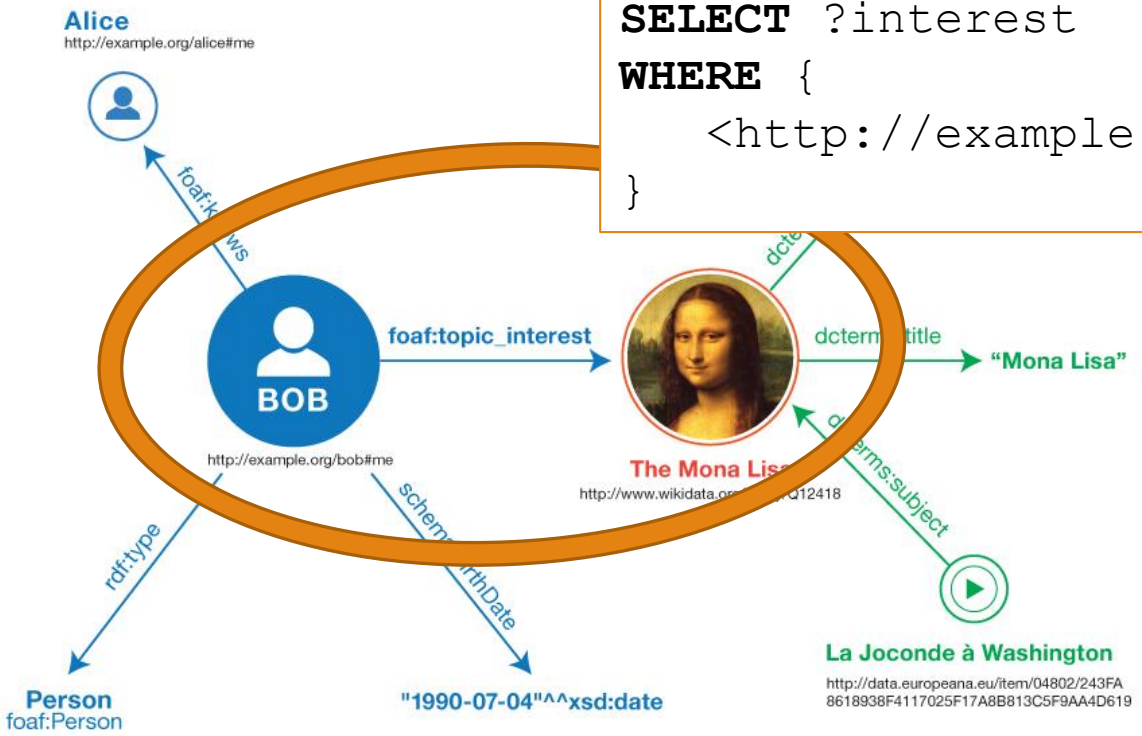
http://example.org/bob#me

http://example.org/alice#me

……

*Why **a** and not **rdf:type**? We are lazy…*

What is bobs interest?



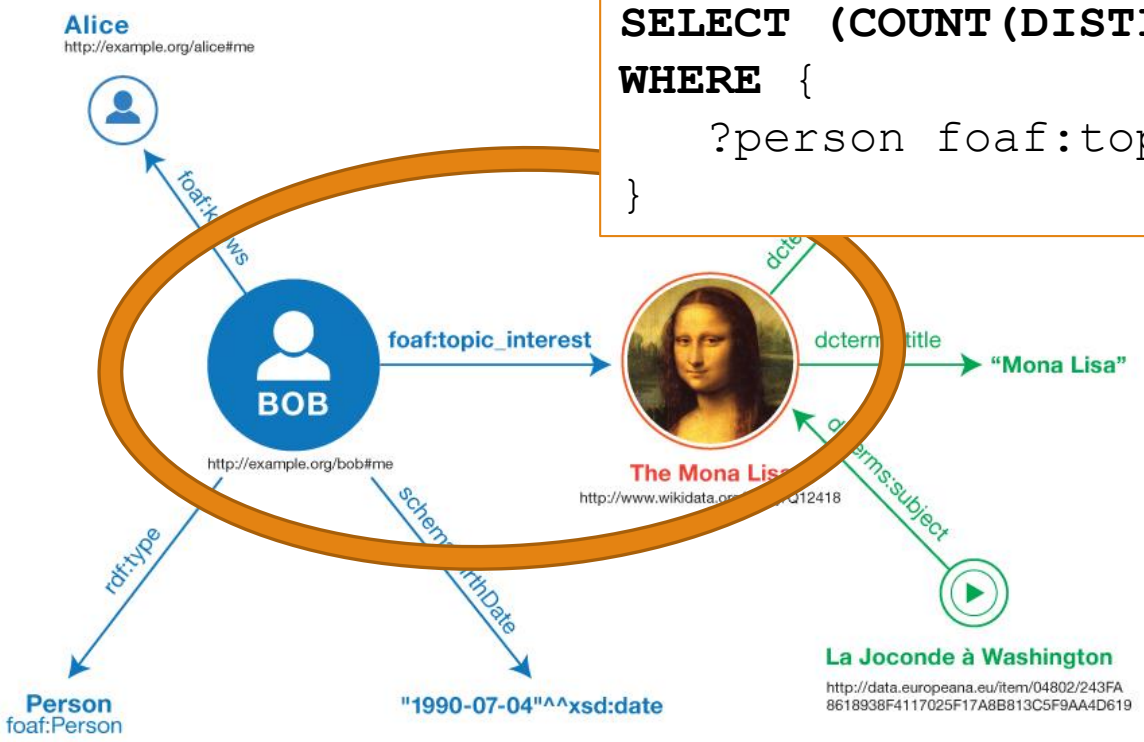

```
PREFIX foaf:<http://xmlns.com/foaf/0.1/>
SELECT ?interest
WHERE {
    <http://example.org/bob#me> foaf:topic_interest ?interest .
}
```

?interest

http://www.wikidata.org/entity/Q12418

If bob was only interested in more topics …

How many people are interested in the mona lisa?



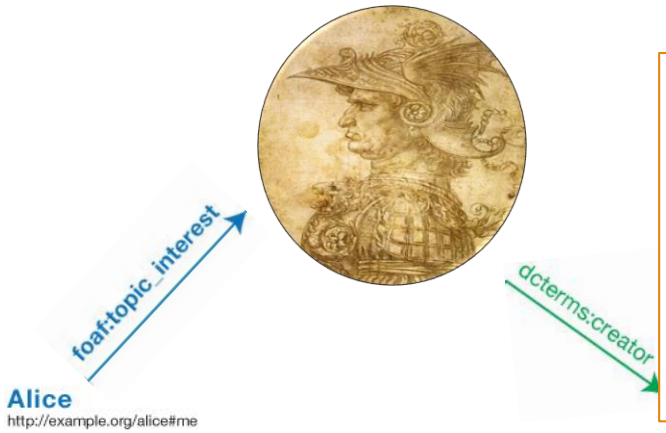

```
PREFIX foaf:<http://xmlns.com/foaf/0.1/>
SELECT (COUNT(DISTINCT(?person)) AS ?pCount)
WHERE {
    ?person foaf:topic_interest <http://www.wikida…/Q12418>.
}
```

?pCount

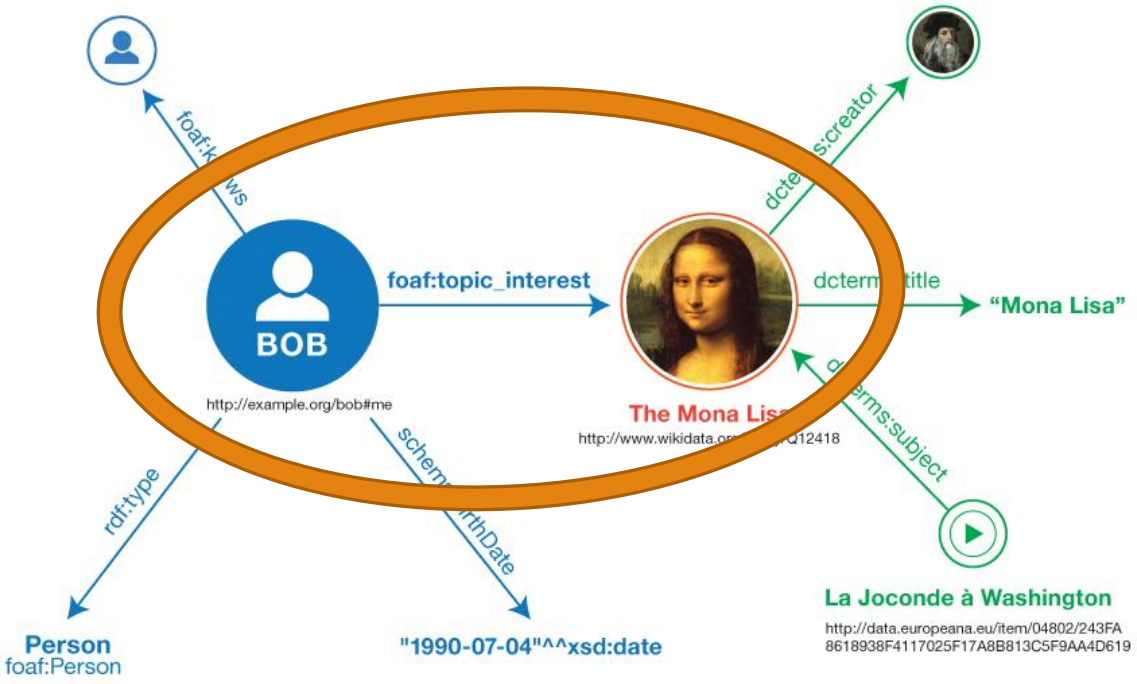
1 (maybe alice likes it but this it not available in the database)

What else is created by and who is interested?:

<http://dbpedia.org/resource/Leonardo_da_Vinci>




```
PREFIX foaf:<http://xmlns.com/foaf/0.1/>
SELECT ?created ?person
WHERE {
    ?created dcterms:creator <http://dbpedia…/Vinci> .
    ?person foaf:topic_interest ?created .
}
```



?created	?person
http://…Q12418	http://…bob#me

… If alice was interested

http://…Q3922633	http://…alice#me

Querying a database

CURRENT "USEFUL" BIOLOGICAL RDF SPARQL ENDPOINTS

http://sparql.uniprot.org/

https://www.ebi.ac.uk/rdf/services/sparql

http://query.wikidata.org/

https://pubchem.ncbi.nlm.nih.gov/rdf/

Your local triple store!

INTERESTING RESOURCES:

Rhea (Reaction database with RDF dump)

BioPax

https://www.bioontology.org/

GBOL.life (current genome annotation ontology)