

# COMP318: Intro on Linked Open Data

[www.csc.liv.ac.uk/~valli/Comp318](http://www.csc.liv.ac.uk/~valli/Comp318)



**Dr Valentina Tamma**

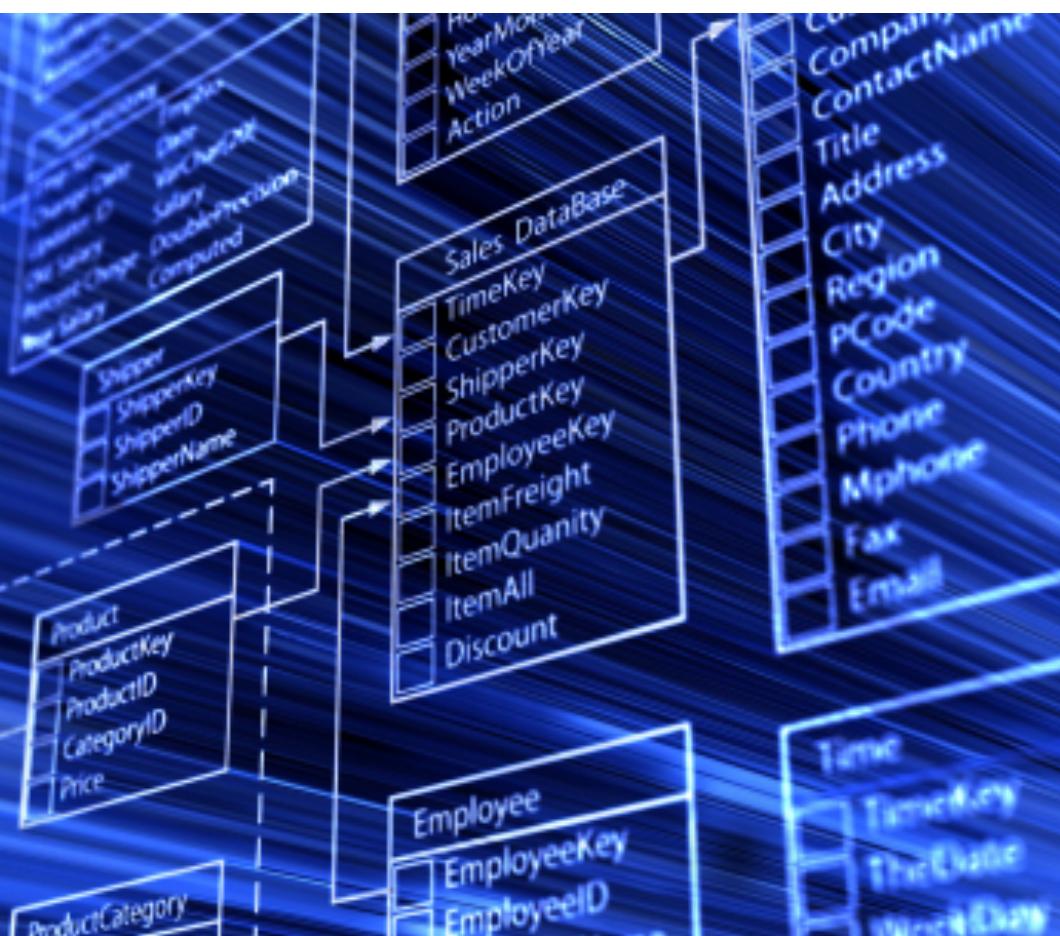
Room: Ashton 2.12

Dept of computer science

University of Liverpool

v.Tamma@liverpool.ac.uk

# The big data shredder



**Structured  
Data**

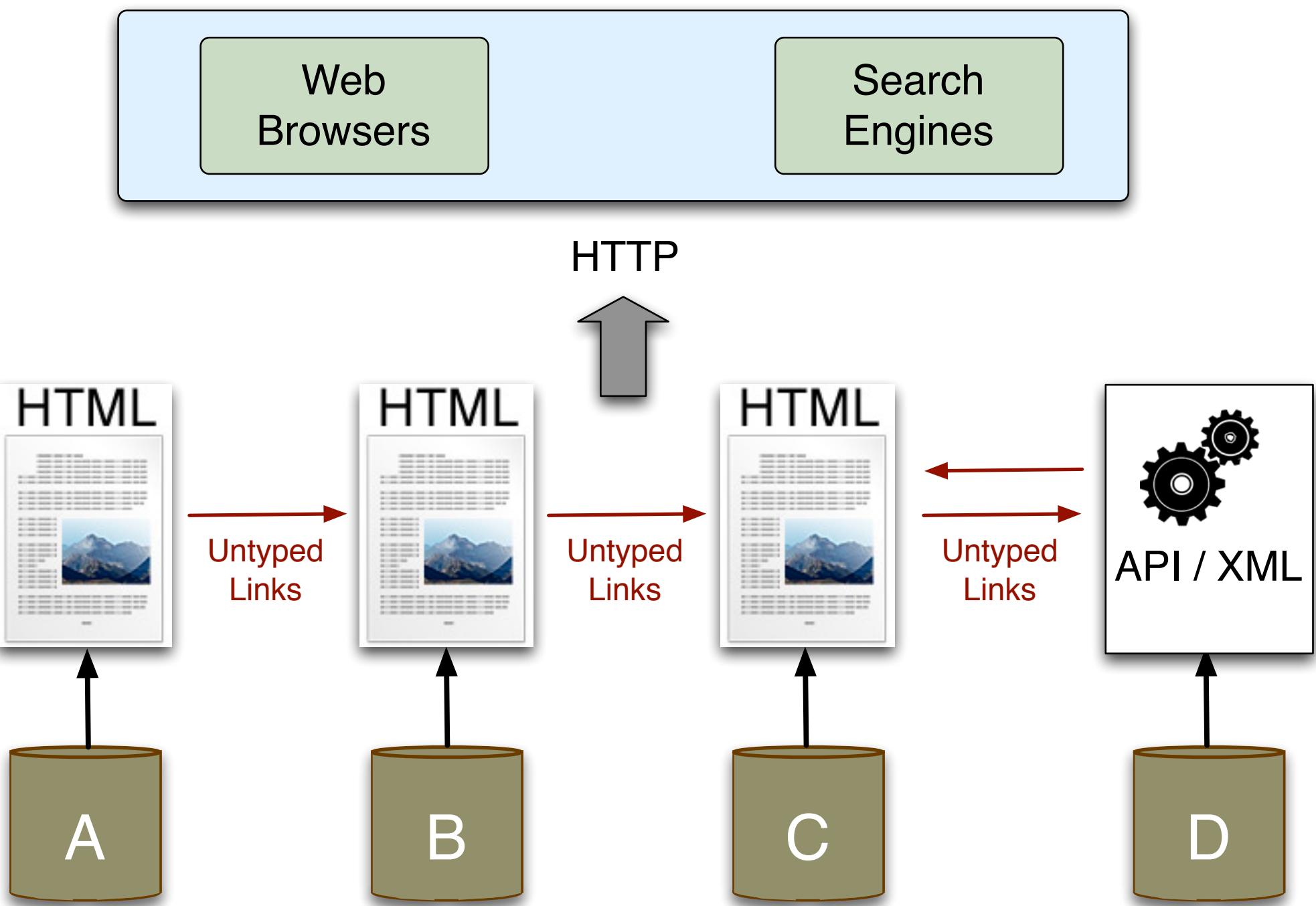


**Unstructure  
Data**

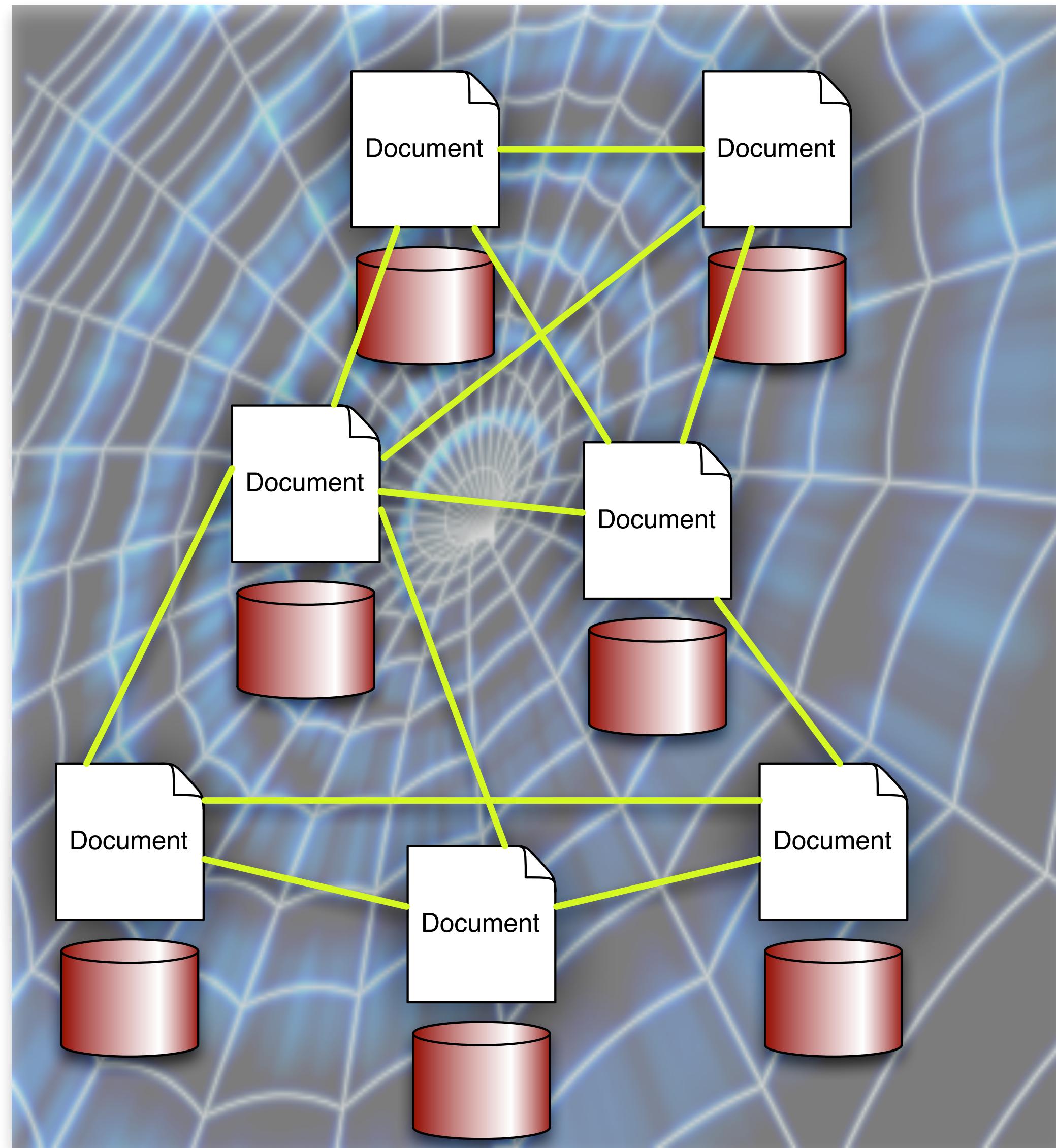
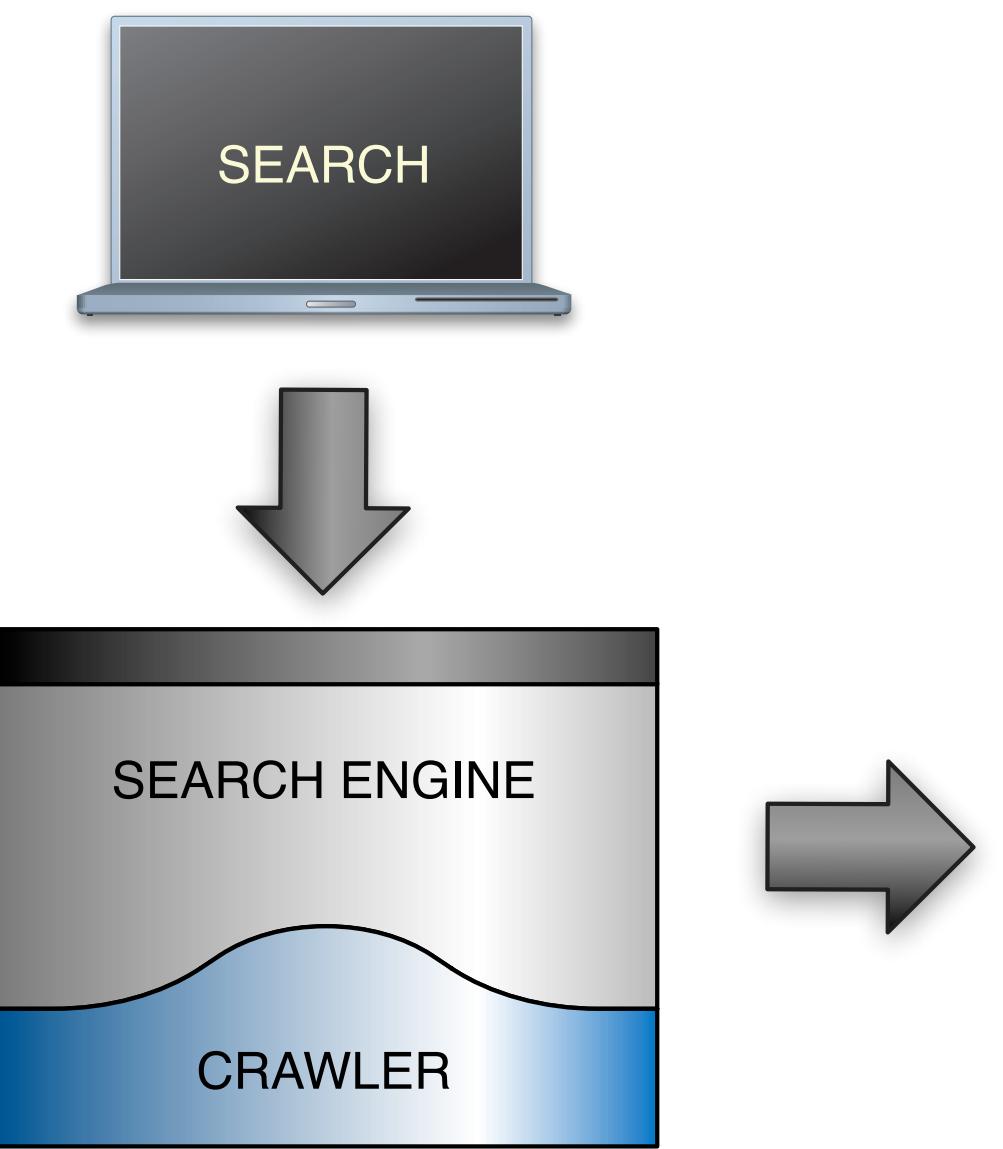
Picture by Martin Hepp

# Back to the beginning: the Web

- The web as a single global information space, similar to a filesystem:
  - Mainly designed for human consumptions;
  - Primary objects: documents;
  - (Hyper) Links between: documents or parts of them;
  - Objects mainly **unstructured** or **semi-structured**;
  - **Implicit** semantics of content and links.

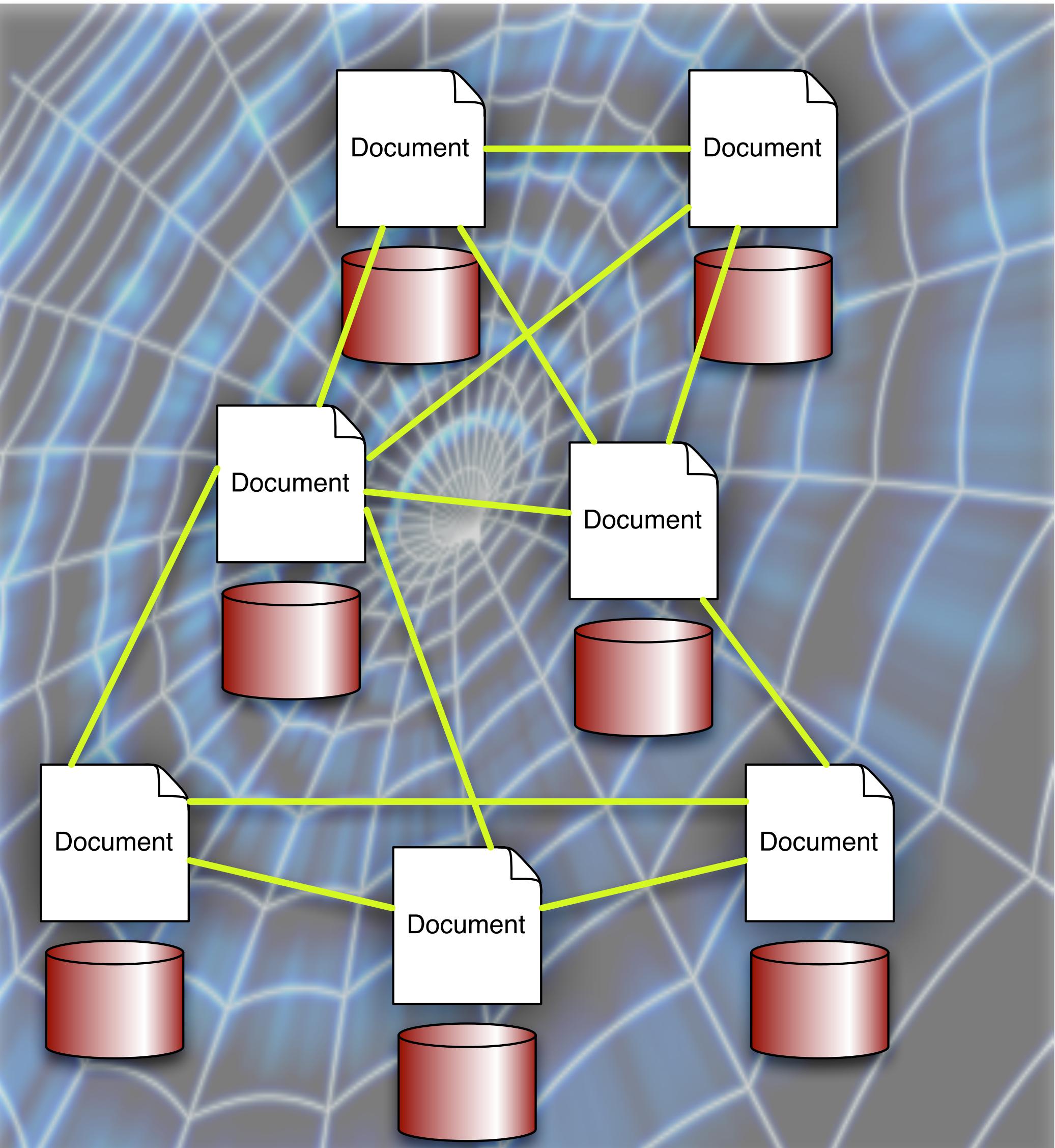


# A web of documents



# A web of documents

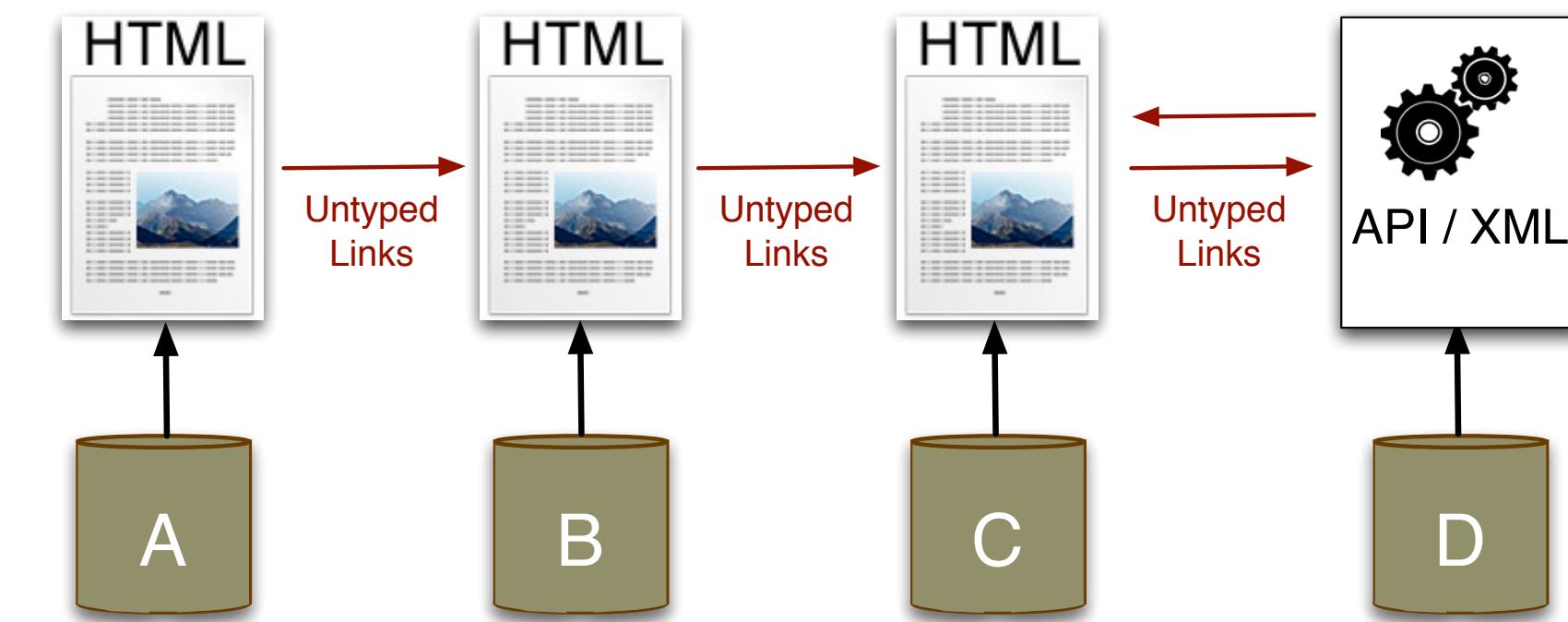
- Is this enough?
  - From a web of documents to
  - ***A web of data***



# Web of Linked Documents

- The problem is that this model does not facilitate “sophisticated” data integration
  - Data integration
    - *Find all the papers published in Semantic Web related conferences in 2011*
  - Querying across different data sources
    - *Find all the papers published in WWW2011 that were authored by people from companies with less than 100 employees.*

?



# Do we search or do we find?

- Ultimate goal is to have software that seamlessly can retrieve the information we need

**Query:** Football players who have played for Manchester City and scored a goal in the 2012 European Championship

Football players who have played for Manchester City and scored a



4,330,000 RESULTS

Narrow by language ▾

Narrow by region ▾

[Manchester City FC History, Players, Team, Information](#)[www.123football.com/clubs/england/manchester-city/index.htm](http://www.123football.com/clubs/england/manchester-city/index.htm) ▾

... of English football in 1999, to being in European ... to spare, City scored a goal to make it 5-1. City ... huge supporters of Manchester City and have played ...

[Ask the Gaffer | Football trivia, questions & answers | Orange UK](#)[web.orange.co.uk/p/web\\_football/ask\\_the\\_gaffer](http://web.orange.co.uk/p/web_football/ask_the_gaffer) ▾

... card and only three yellows - and he's scored a goal. ... Who are 9 players who have played for two of the Big Four (ManU ... foreign players Man City ...

[Archive \(part two\) | Football | guardian.co.uk](#)[www.guardian.co.uk/football/2005/sep/21/theknowledge.sport](http://www.guardian.co.uk/football/2005/sep/21/theknowledge.sport) ▾

21/09/2005 · ... European leagues? Players who've played for two clubs twice; Which football ... Manchester City; current Premiership players to have played ... Players who have scored ...

[Football players profiles - English footballers soccer careers](#)[www.football-england.com/football\\_players\\_profiles.html](http://www.football-england.com/football_players_profiles.html) ▾

Football players profiles - Accounts ... forever, wasn't he Scored a ponfull of goals. John Gidman Solid full back played for Aston Villa, Man Utd ... Francis Lee Roly-poly Man City ...

[Club Profile | Manchester City](#)[www.premierleague.com/page/manchester-city-fc](http://www.premierleague.com/page/manchester-city-fc) ▾

We have updated our cookie policy to reflect ... Last six results; BPL: 9 Feb 2013: v Southampton (A) 3 - 1 ... Official Manchester City Football Club website

[BBC Sport - Football - Championship](#)[www.bbc.co.uk/sport/0/football/championship](http://www.bbc.co.uk/sport/0/football/championship) ▾

11/04/2013 · The latest BBC Championship news plus live scores, fixtures ... Watch video Leicester man hoping for goal boost ... position, team name, games played, total goal ...

[Manchester derby - Wikipedia, the free encyclopedia](#)[en.wikipedia.org/wiki/Manchester\\_derby](http://en.wikipedia.org/wiki/Manchester_derby) ▾

History · Statistics · Honours · Non-competitive derbies · All-time results

... football supporters in Manchester watched City ... City in the 24th and 45th minutes respectively, and Michael Carrick scored a consolation goal for ... Players who have played ...

[Arsenal FC vs Manchester City LIVE Commentary - Goal.com](#)[www.goal.com/en-gb/match/60550/arsenal-fc-vs-manchester-city/play...](http://www.goal.com/en-gb/match/60550/arsenal-fc-vs-manchester-city/play...) ▾

... of football ... has scored a goal every 94 minutes this season, the best rate in the Premier League for all players who have played ... his Manchester City players. Having played ...

[Ask the Gaffer | Football trivia, questions & answers | Orange UK](#)[web.orange.co.uk/p/football/ask\\_the\\_gaffer](http://web.orange.co.uk/p/football/ask_the_gaffer) ▾

Web Images Maps Shopping More Search tools

About 27,700,000 results (0.41 seconds)

[Football news, results, live scores, rumours and fixtures - Goal.com](#)[www.goal.com/en-gb/](http://www.goal.com/en-gb/)

Get the latest Football news, scores, results, rumours, fixtures and live action from the premier ... had earlier scored the vital second goal to send Manchester City into the FA Cup final ... Player Ratings: Chelsea 1-2 Manchester City » ... Europe's premier club competition returns as the knockout stages begin and there are ...

[Man City | Manchester City | Football Club News, Results ... - Goal...](#)[www.goal.com/en/teams/england/109/manchester-city-fc](http://www.goal.com/en/teams/england/109/manchester-city-fc)

Europe Home · Europe News · England Home · England Table/Results · Italy Home ... COMMENT: Brilliant Bayern show why the Champions League is theirs to lose » ...

PLAYER RATINGS: Manchester United 1-2 Manchester City » ... and admits he has never watched a repeat of his side's famous defeat from last season ...

[Football News, Football Scores, Football Transfers ... - Goal.com](#)[www.goal.com/en/](http://www.goal.com/en/)

Get the latest worldwide football news, football scores, transfer rumours, football ... The Portuguese star scored the second-fastest goal of his career before firing in ...

PLAYER RATINGS: Chelsea 1-2 Manchester City » .... The Bayern Munich defender has urged his team-mates to approach the Champions League tie against ...

[News for Football players who have played for ...](#)[Chelsea left playing catch-up as Manchester City superiority tells](#)[The Guardian \(blog\)](#) - 8 hours ago

Chelsea left playing catch-up as Manchester City superiority tells ... as he tumbled to the turf to score the first goal conceded by Manchester City in this ... The European champions have experienced it a few times during a season ... Premier League last year and are seven points apart in this season's table.

[Manchester City F.C. - Wikipedia, the free encyclopedia](#)[en.wikipedia.org/wiki/Manchester\\_City\\_F.C.](http://en.wikipedia.org/wiki/Manchester_City_F.C.)

Manchester City Football Club is an English Premier League football club based ... The club has played at the City of Manchester Stadium since 2003, having ... League Championship, FA Cup, League Cup and European Cup Winners' ... the following season, despite scoring more goals than any other team in the division.

[Premier League - Wikipedia, the free encyclopedia](#)[en.wikipedia.org/wiki/Premier\\_League](http://en.wikipedia.org/wiki/Premier_League)

At the top of the English football league system, it is the country's primary football ... The current champions are Manchester City, who won the title in the 2011–12 season. ... clubs playing in European competitions in 1990 (resulting in Manchester ... The first ever Premier League goal was scored by Brian Deane of Sheffield ...

# What we were actually looking for

Create account Log in

Article Talk Read View source View history Search

**Mario Balotelli**

From Wikipedia, the free encyclopedia

**Mario Barwuah Balotelli** (Italian pronunciation: [maˈrjo baloˈtelli]; born Mario Barwuah; 12 August 1990) is an Italian footballer who plays as a striker for Milan and the Italian national team.<sup>[4][5]</sup> He started his professional football career at Lumezzane and played for the first team twice before having an unsuccessful trial at FC Barcelona,<sup>[6]</sup> and subsequently joining Internazionale in 2007. Inter manager Roberto Mancini brought Balotelli into the first team, but when Mancini left, Balotelli's disciplinary record fell away. He had a strained relationship with new head coach José Mourinho and was suspended from Inter's first team in January 2009 after a number of disciplinary problems.

In March 2010, he came under criticism by Inter fans after he appeared on the Italian TV show *Striscia la notizia*, wearing an Milan jersey. This damaged the prospect of him having a long career at Inter, but he did make several appearances after that. With doubts over his career at Inter, former coach Roberto Mancini had since moved to Manchester City and decided to give Balotelli a fresh chance at a new club. He joined Manchester City in August 2010, where his performances and off-field activities have continued to be enigmatic and unpredictable. Nicknamed Super Mario,<sup>[7]</sup> he earned his first cap for the Italian national team on 10 August 2010 in a friendly match against the Côte d'Ivoire.

Contents [hide]

- 1 Early life
- 2 Club career
  - 2.1 Lumezzane
  - 2.2 Internazionale
  - 2.3 Manchester City
    - 2.3.1 2010–11
    - 2.3.2 2011–12
    - 2.3.3 2012–13
  - 2.4 Milan
- 3 International career
  - 3.1 UEFA Euro 2012
  - 3.2 2014 FIFA World Cup qualifying
  - 3.3 International goals
- 4 Style of play
- 5 Personality and reputation
- 6 Career statistics
  - 6.1 Club
  - 6.2 International
- 7 Honours
  - 7.1 Club
  - 7.2 International
  - 7.3 Individual
- 8 Outside football
  - 8.1 Personal life
- 9 References
- 10 External links

**Early life**

**Mario Balotelli**



Balotelli playing for Italy at the UEFA Euro 2012

**Personal information**

Full name	Mario Barwuah Balotelli <sup>[1]</sup>
Date of birth	12 August 1990 (age 22) <sup>[2]</sup>
Place of birth	Palermo, Italy
Height	1.89 m (6 ft 2 in) <sup>[3]</sup>
Playing position	Striker <sup>[2]</sup> / Forward

**Club information**

Current club	Milan	
Number	45	
Youth career		
2001–2006	Lumezzane	
Senior career*		
Years	Team	Appst (Gls)†
2005–2006	Lumezzane	2 (0)
2006–2010	Internazionale	59 (20)
2010–2013	Manchester City	54 (20)
2013–	Milan	7 (7)

**National team**

Was there any other player???

# Why can't just we find it



BIBTEX



# Mashups?

- **Data:** the value of observable, measurable or calculable attributes;

Data

- **Information:** is data plus conceptual commitments and interpretations (*context*) that provides some meaning.

- Information is data extracted, filtered or formatted in some way
- the context includes the class to which the attributes belong, the objects that are members of that class, object operations or behaviour, and the relationships to other objects and classes

Mashups

- **Knowledge:** an elusive concept. information having been processed, organised or structured in some way, or else as being applied or put into action.

- Knowledge as synthesis of information
- Knowledge as know how
- Knowledge as propositional (beliefs)

Information

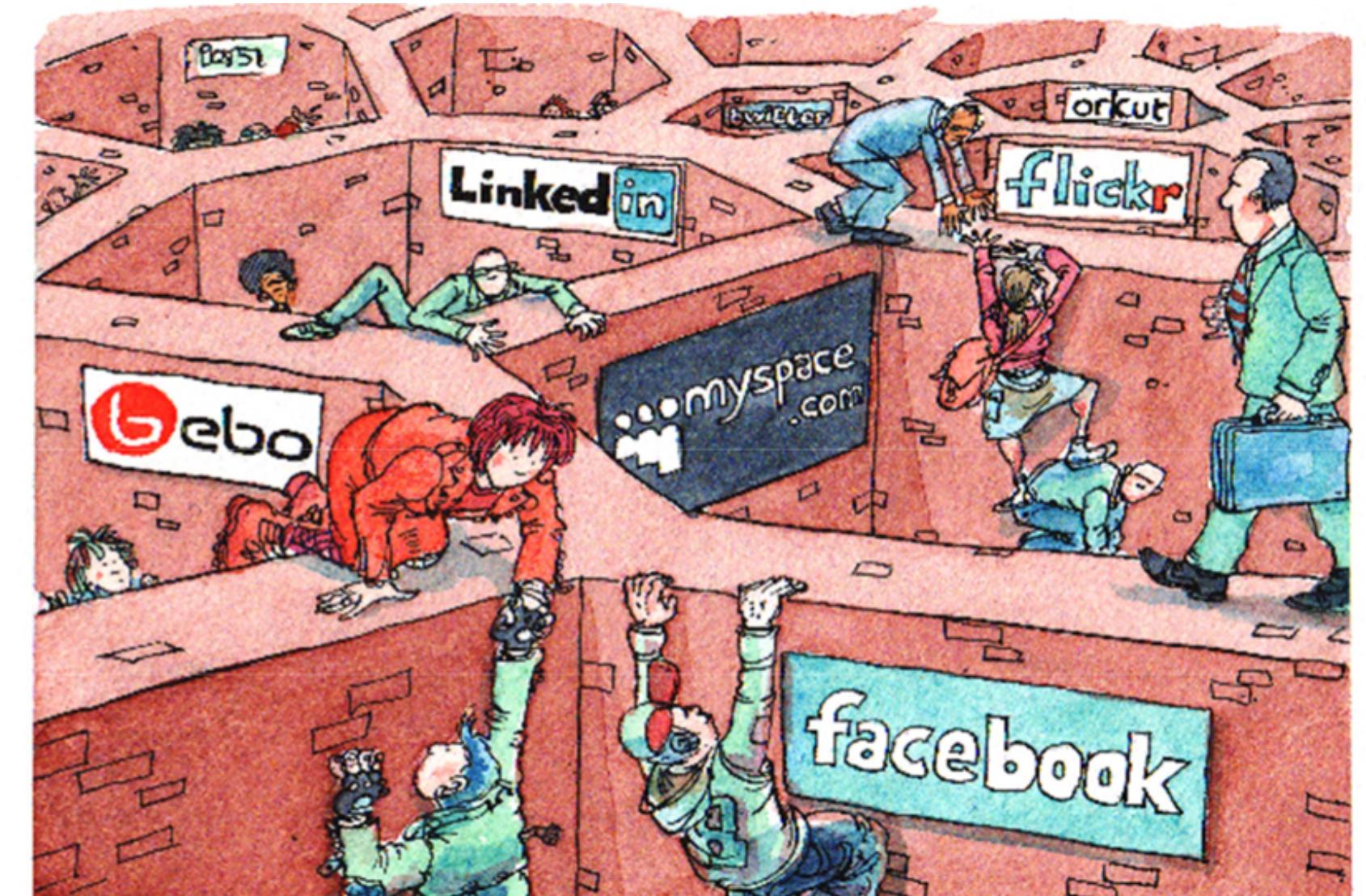
Knowledge

# Mashups?

- Mashups are web pages or applications that utilise and combine data, presentation or functionalities from several data sources in order to create new services
  - *Use APIs provided by different content sites to aggregate and reuse the content in different ways*
  - *Combination, aggregation, visualisation...*
- Different types of mashups
  - **Business (or enterprise) mashups** define applications that combine their own resources, application and data, with other external Web services.
    - They are secure, visually rich Web applications that expose actionable information from diverse internal and external information sources.
  - **Consumer mashups** combines different data types. It combines data from multiple public sources in the browser and organises it through a simple browser user interface
    - Wikipediavision combines Google Map and a Wikipedia API
  - **Data mashups** combine similar types of media and information from multiple sources into a single representation. The combination of all these resources create a new and distinct Web service that was not originally provided by either source.

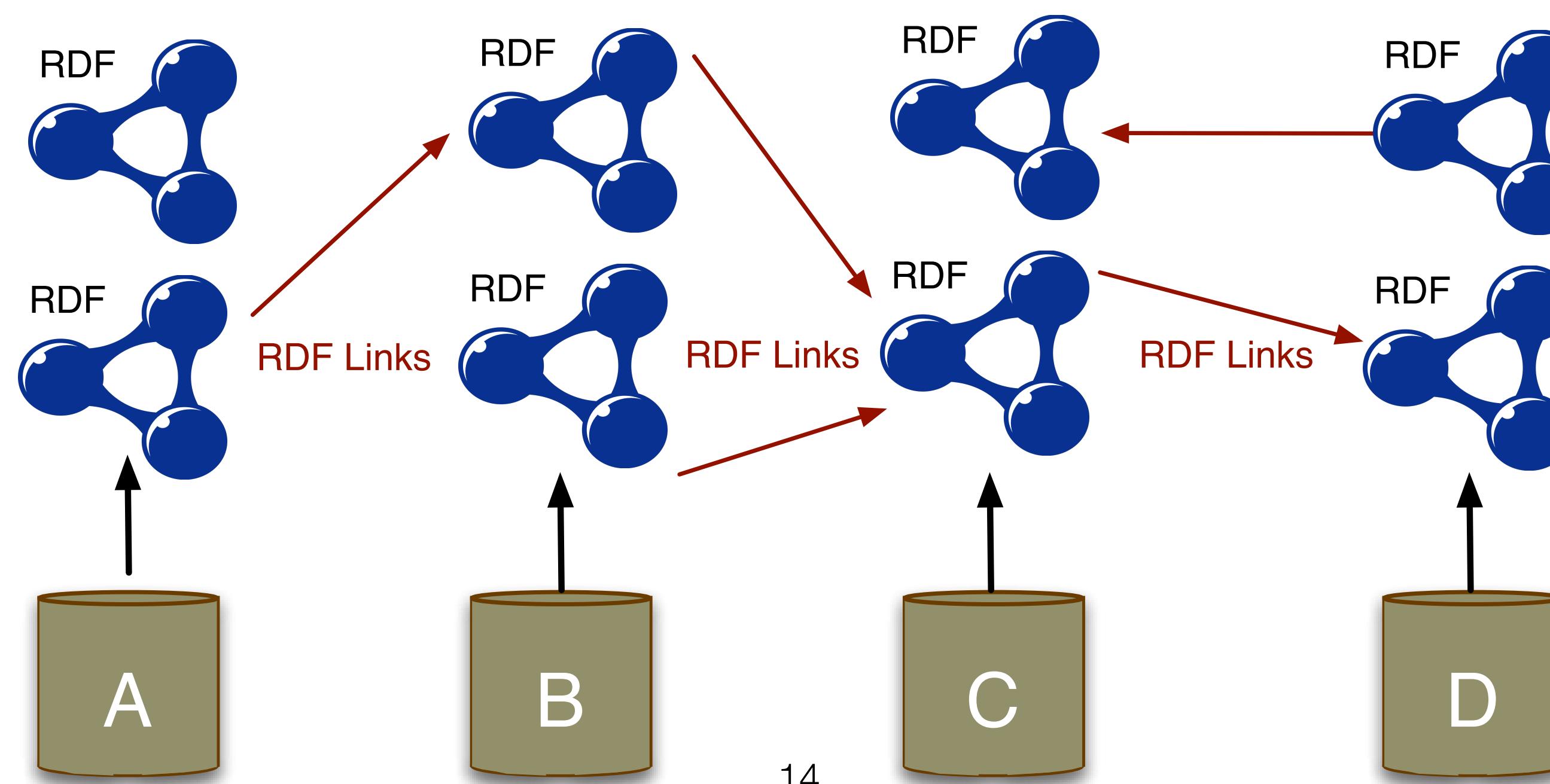
# Problem with mashups and API

- API allow the mashing up of information but...
  - APIs have proprietary interfaces
  - Mashups are based on a fixed set of data sources
  - There are no **(hyper)links** that relate data items within different APIs
    - the data is disconnected
    - these are all information and data siloes



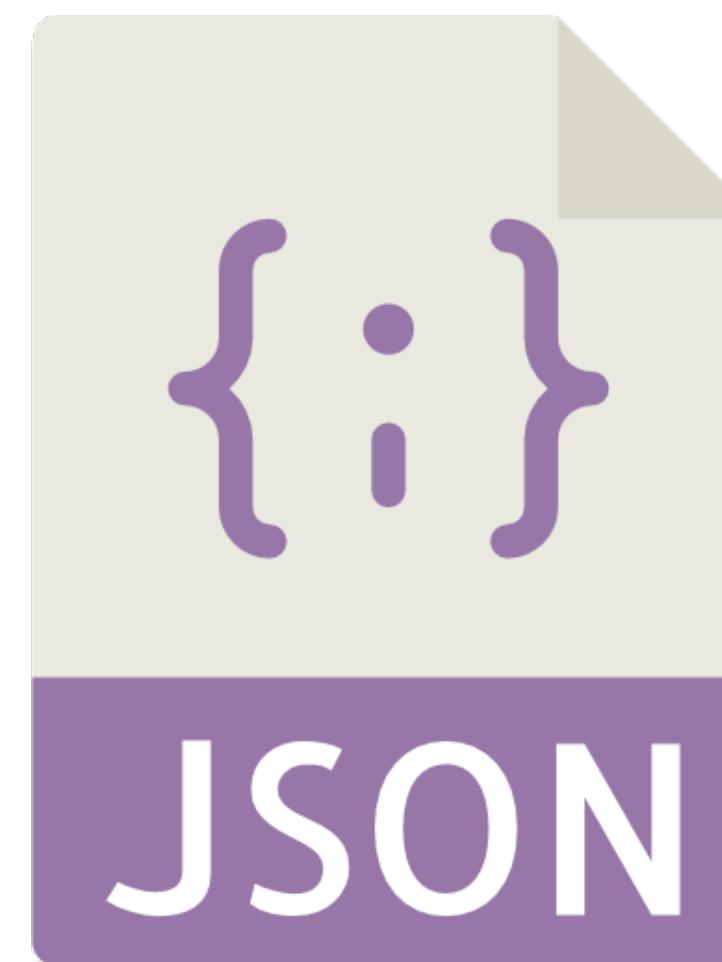
# Linked data: the idea

- Use the Web to create a global dataspace
  - RDF as a means to publish structured data on the Web
  - state explicitly links between data items from different data sources
  - Link data rather than documents!



# Web of data

- Isn't this what we are doing now?
  - CSV, XML, key-value data, relational data
- But it's about the *semantics* rather than the surface or exchange syntax



# In Other Words

- “All kinds of conceptual things, they have names now that start with HTTP.”
- “I get important information back. I will get back some data in a standard format which is kind of useful data that somebody might like to know about that thing, about that event.”
- “I get back that information it's not just got somebody's height and weight and when they were born, it's got relationships. And when it has relationships, whenever it expresses a relationship then the other thing that it's related to is given one of those names that starts with HTTP.”

It is about:  
a data (RDF)  
and naming (URI)  
model on Web

*Tim Berners-Lee: Linked Open Data Presentation, TED 2009*

# What is linked data

- Linked Data refers to a set of best practices for publishing and interlinking structured data on the Web:
  - Based on 4 principles, following the web architecture:
    1. Use URIs as to identify things;
    2. Use HTTP URIs so that these things can be referred to and looked up ("dereferenced") by people and user agents.
    3. Provide useful information about the thing when its URI is dereferenced, using standard formats such as RDF/XML.
    4. Include links to related URIs in the exposed data to improve discovery of other related information on the Web.
      1. (a bit obvious, maybe) Link the data!



Tim Berners-Lee: <http://www.w3.org/DesignIssues/LinkedData>

# 1. Use URIs as to identify things

- Use URIs for naming everything you could think of:
  - URI references should be used to identify not just Web documents and digital content, but also:
  - real world objects and abstract concepts:
    - tangible things
      - e.g. people, places and cars,
    - more abstract things
      - e.g. the relationship type of knowing somebody, the set of all green cars in the world, or the colour green itself
- This principle extends the scope of the Web:
  - from online resources to encompass any object or concept in the world.

## 2. Use HTTP URIs allowing to refer and look up things on the Web

- HTTP protocol as the Web's universal access mechanism for people and user agents to dereference resources.
  - In the classic Web, HTTP URIs are used to combine globally unique identification with a simple, well-understood retrieval mechanism.
  - Thus, Linked Data advocates the use of HTTP URIs to identify objects and abstract concepts:
    - Dereferenceable URIs are recommended but not mandatory in linked data. When you do an HTTP GET (or type in a URL in your browser's address bar) something is returned, an RDF description in the best case!
    - Different servers are responsible for answering requests attempting to dereference HTTP URIs in many different namespaces.

### 3. Assign useful information to URIs using standards

- Provide useful information about the thing when its URI is dereferenced,
  - using standard formats such as RDF/XML or OWL
  - agreement on standardised content formats!
- Advocates the use of a single data model for publishing structured data on the Web
  - Resource Description Framework (RDF)
  - simple graph-based data model.

# 4. Include links to other URLs to improve discovery

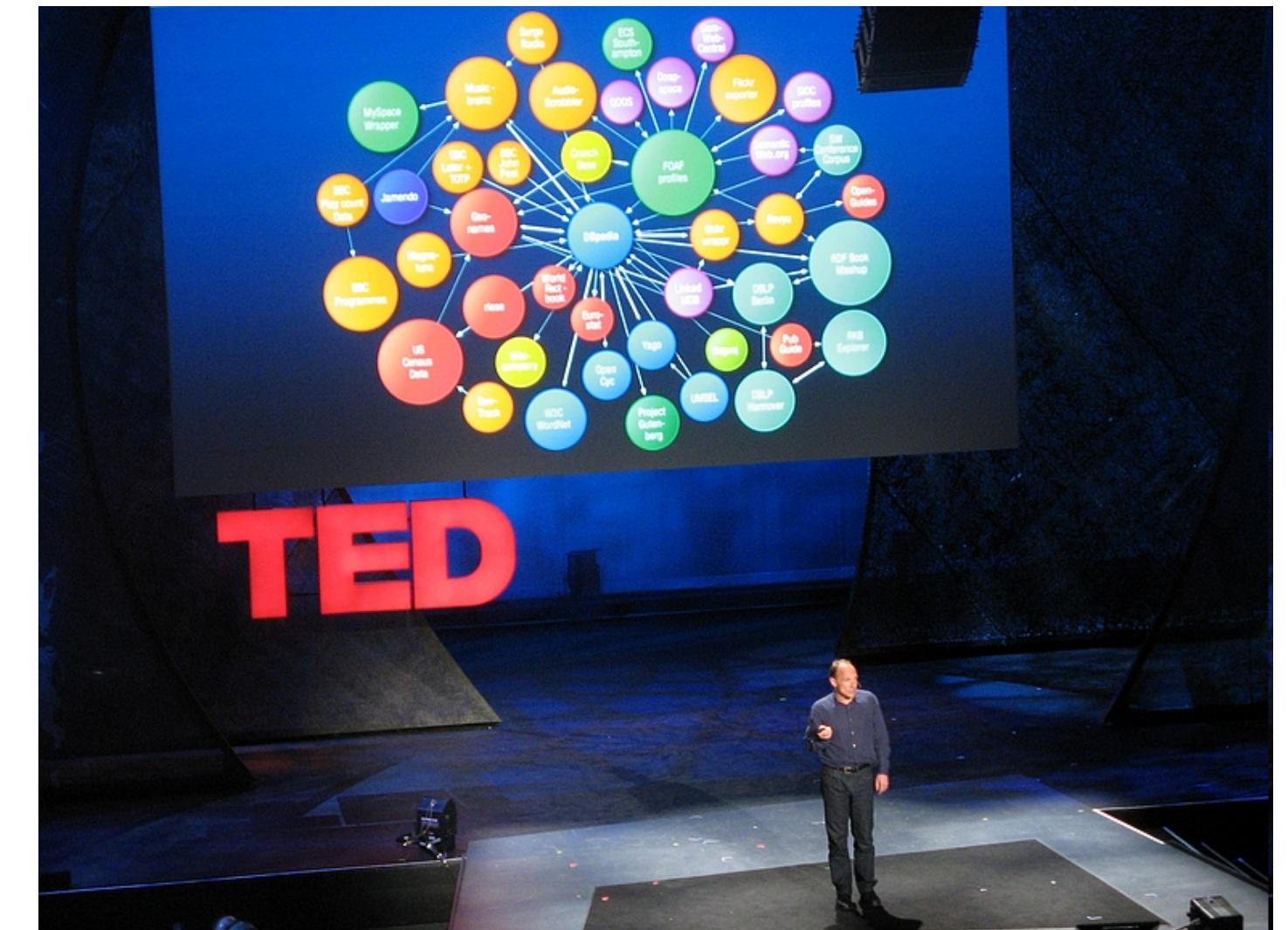
- Following from principle 1, hyperlinks connect not only Web documents, but any type of thing.
  - e.g. a hyperlink may be set between a person and a place, or between a place and a company.
- Hyperlinks that connect things in Linked Data have types which describe the relationship between the things.
  - Unlike hyperlinks on the Web.
  - e.g. a hyperlink of the type friend of may be set between two people, or a hyperlink of the type based near may be set between a person and a place.
  - When an RDF link connects URLs in different namespaces, it ultimately connects resources in different data sets
    - **These typed hyperlinks are RDF properties interpreted as hyperlinks**

# 4a. Link the data!

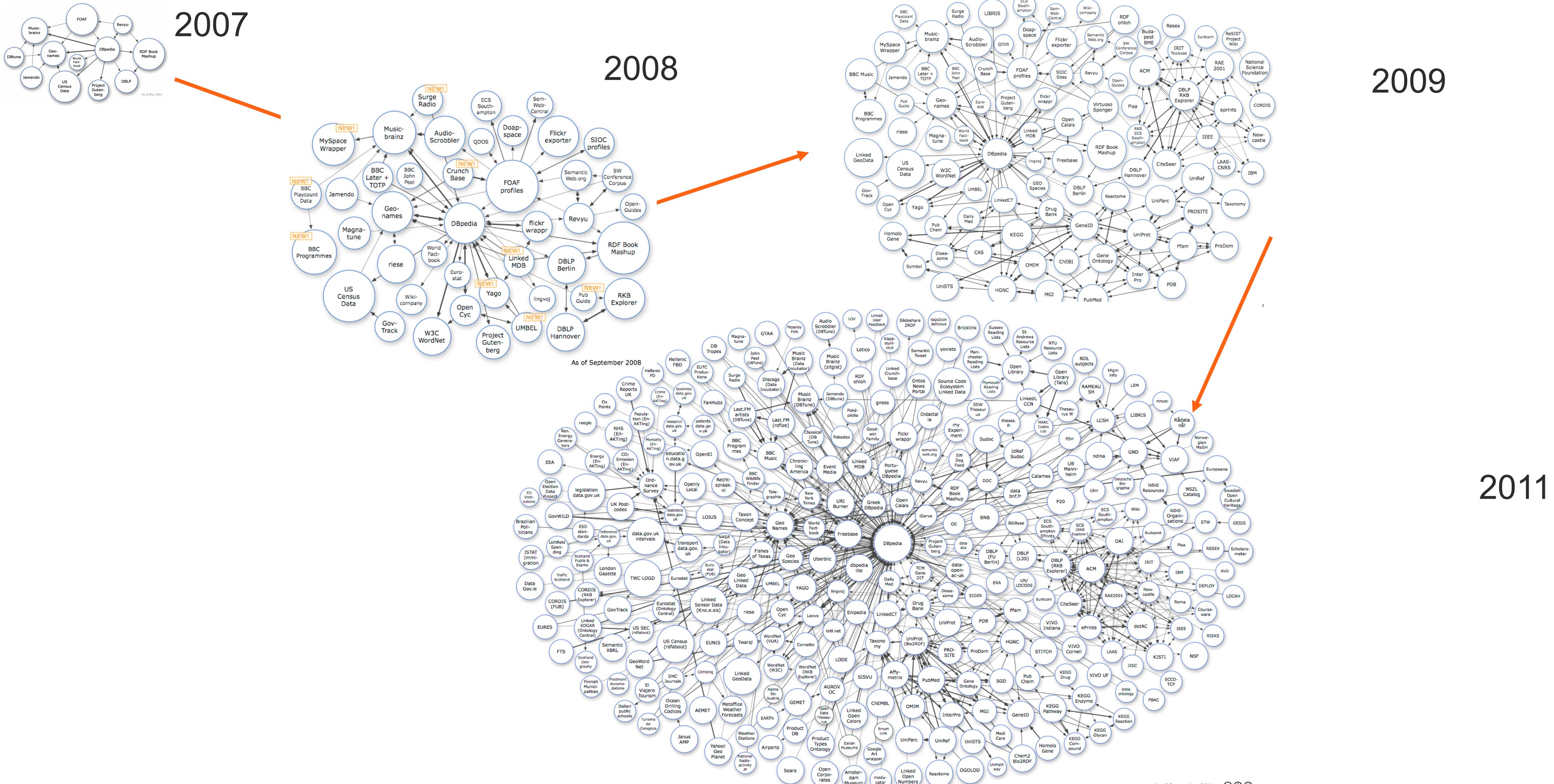
- Across the Web, many different servers are responsible for answering requests attempting to dereference HTTP URIs in many different namespaces:
  - In Linked Data, this means returning RDF descriptions of the resources identified by these URIs.
    - If an RDF link connects URIs in different namespaces, it ultimately connects resources in different data sets.
- Linked Data uses hyperlinks to connect disparate data into a single global data space.
  - just as hyperlinks in the classic Web connect documents into a single global information space,
- These links, in turn, enable applications to navigate the data space.
  - e.g. a Linked Data application that has looked up a URI and retrieved RDF data describing a person may follow links from that data to data on different Web servers, describing the place where the person lives or the company for which the person works.

# But TBL was talking about open data i.e. when the sw meets open data

- Over the last few years, the Linked Data paradigm has found a huge application when combined with the publication of data with liberal licenses.
- The Open Data Movement, aims to release huge data sets often from local government authorities.
- It uses Linked Data technologies and best practices to publish a plethora of different interlinked data sets.
  - a bunch of data published with an open license is intended to be freely available to everyone to use and republished without restriction from copyright, patents or other restrictions.
- Where Open Data meets Linked Data, we have Linked Open Data.



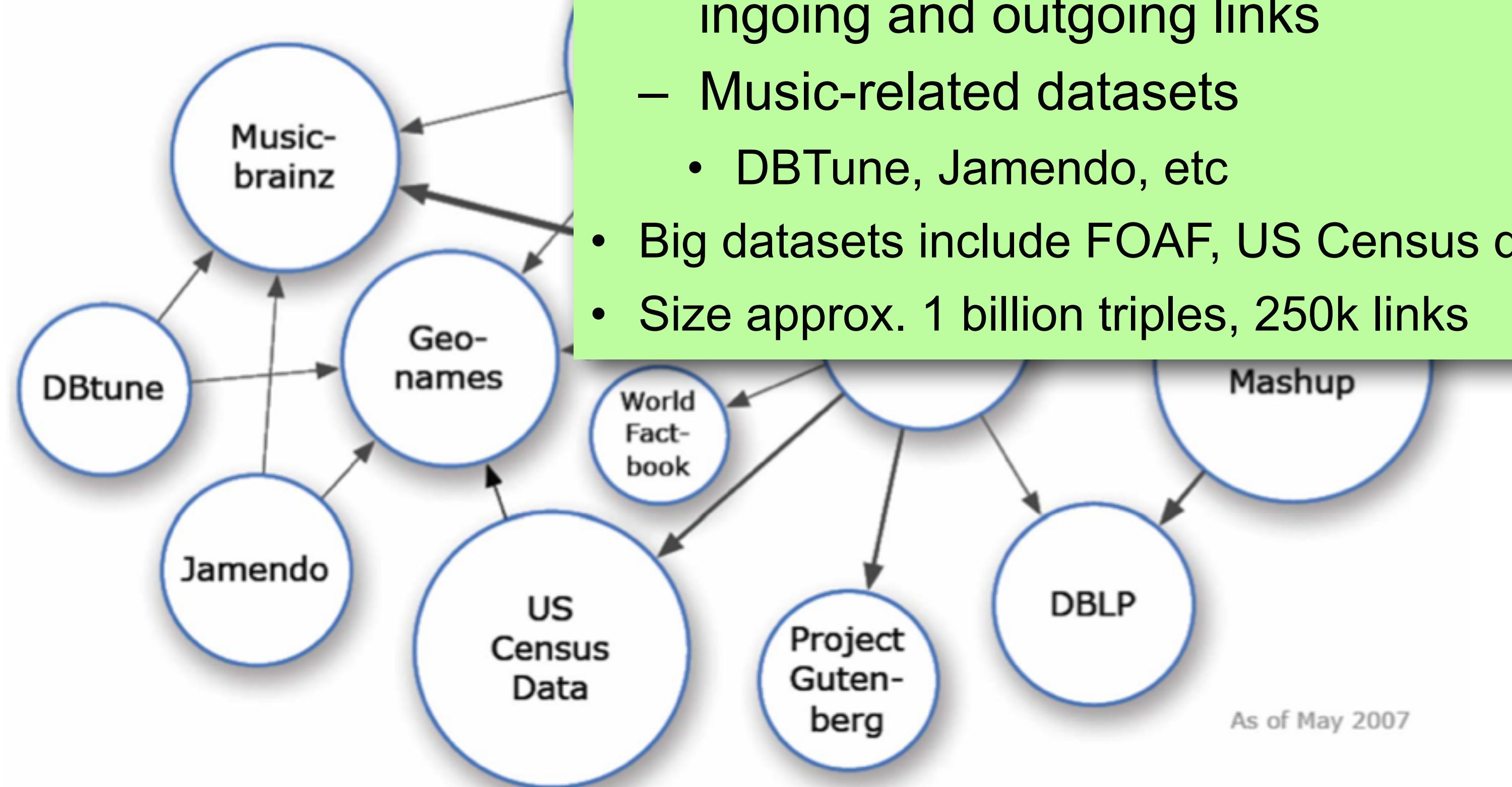
# Linked Data Evolution: is there data out there?



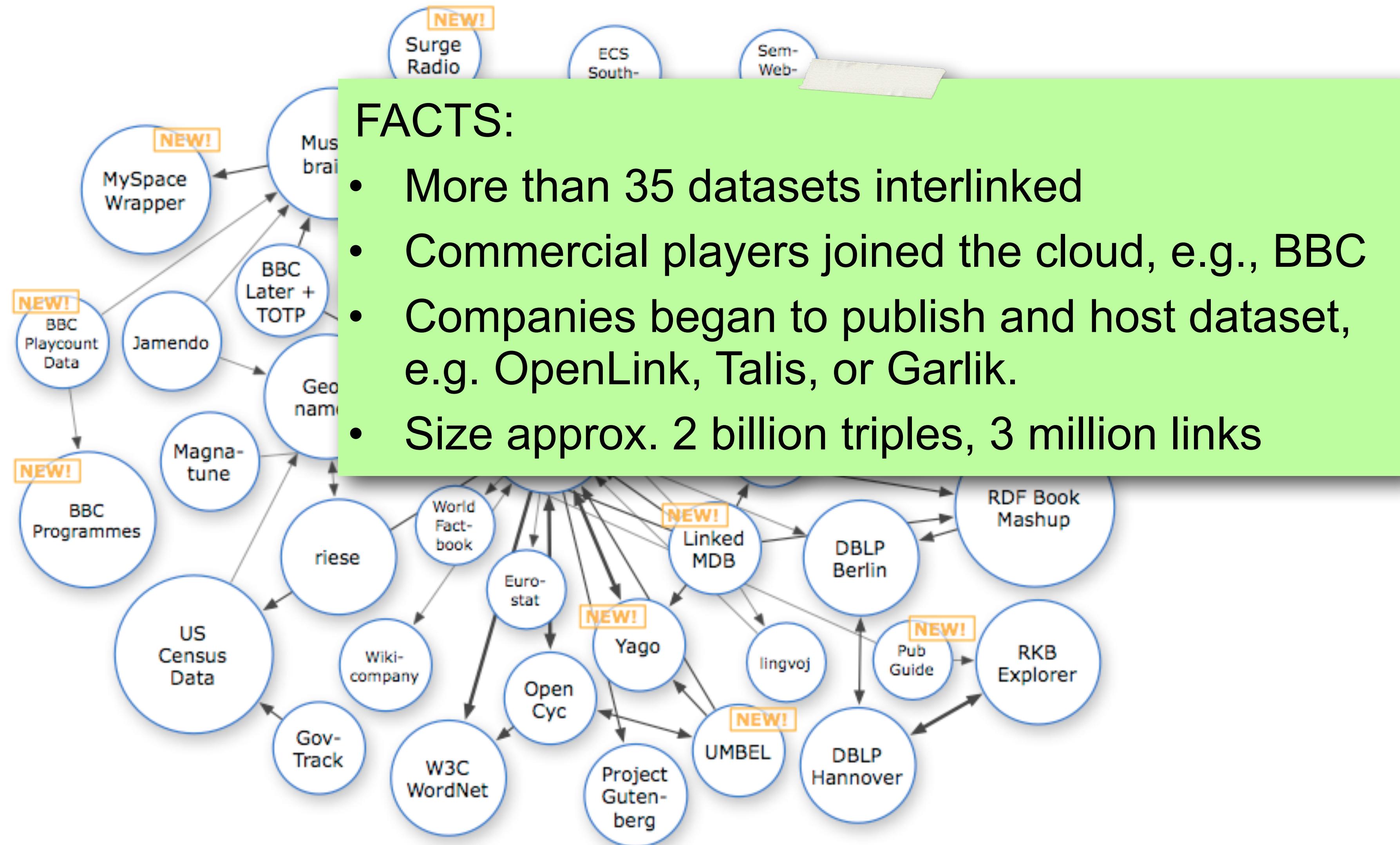
# May 2007

## FACTS:

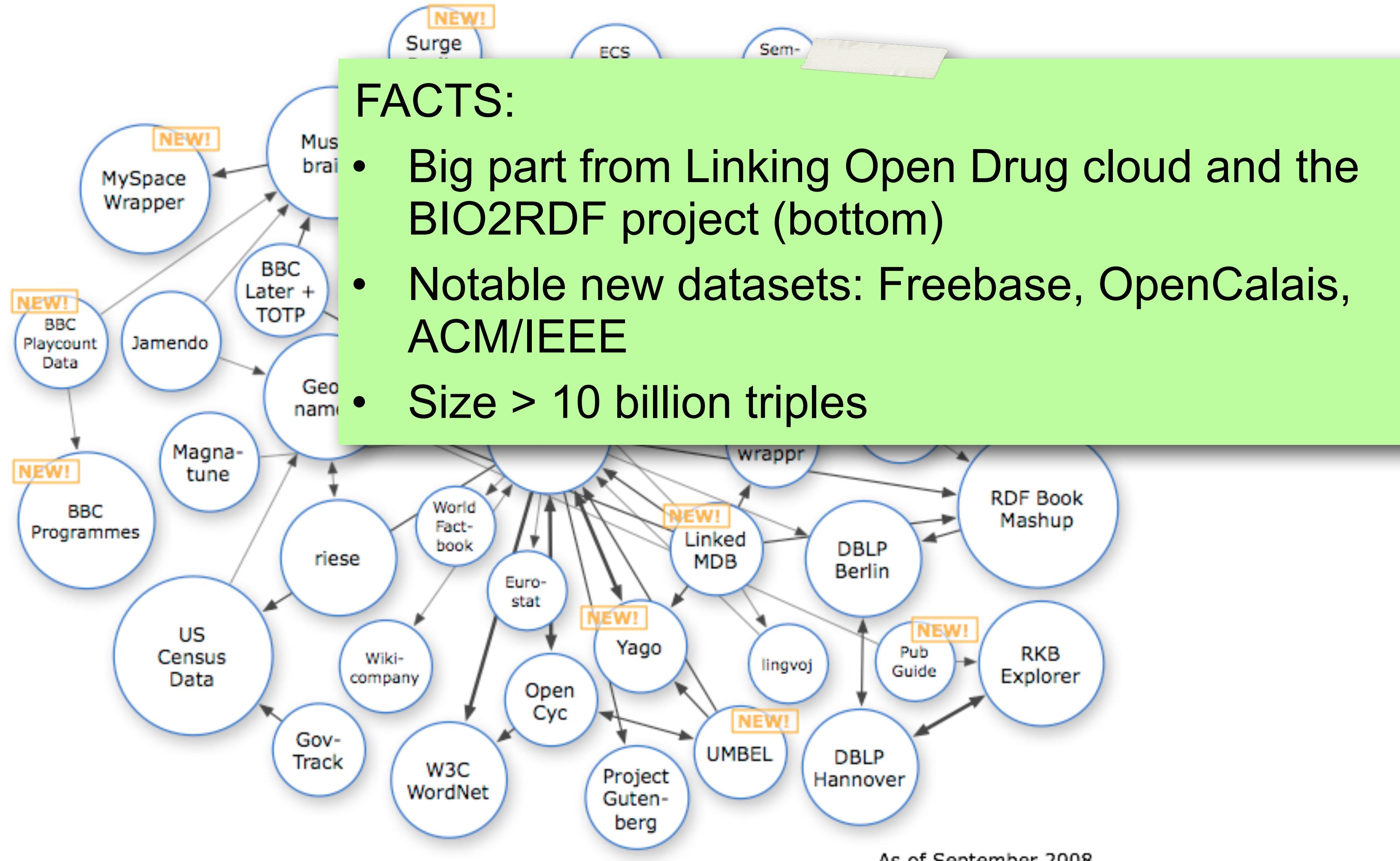
- Focal point based
  - DBpedia: RDFized version of Wikipedia; many ingoing and outgoing links
  - Music-related datasets
    - DBTune, Jamendo, etc
- Big datasets include FOAF, US Census data
- Size approx. 1 billion triples, 250k links



# September 2008

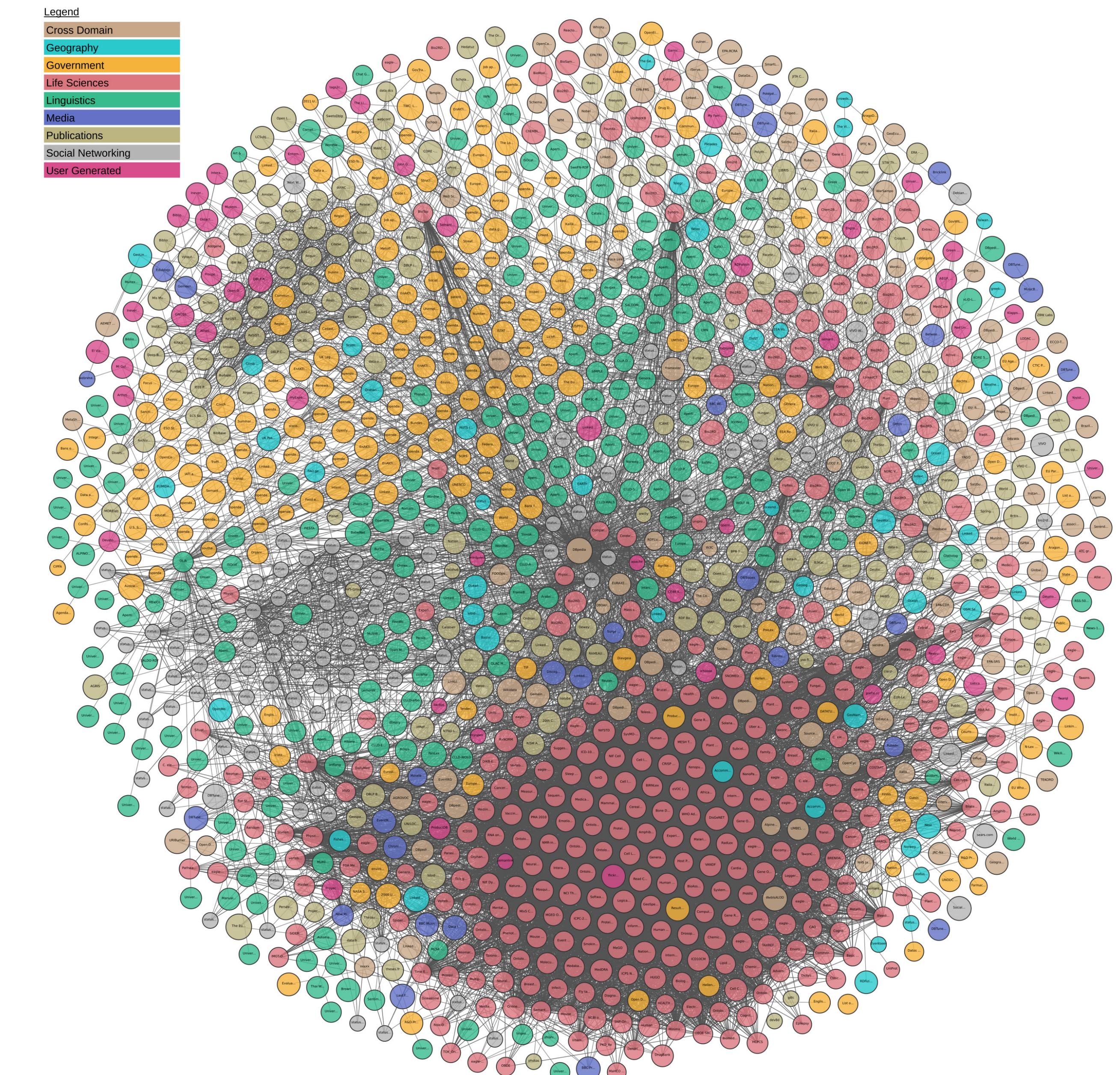


# March 2009

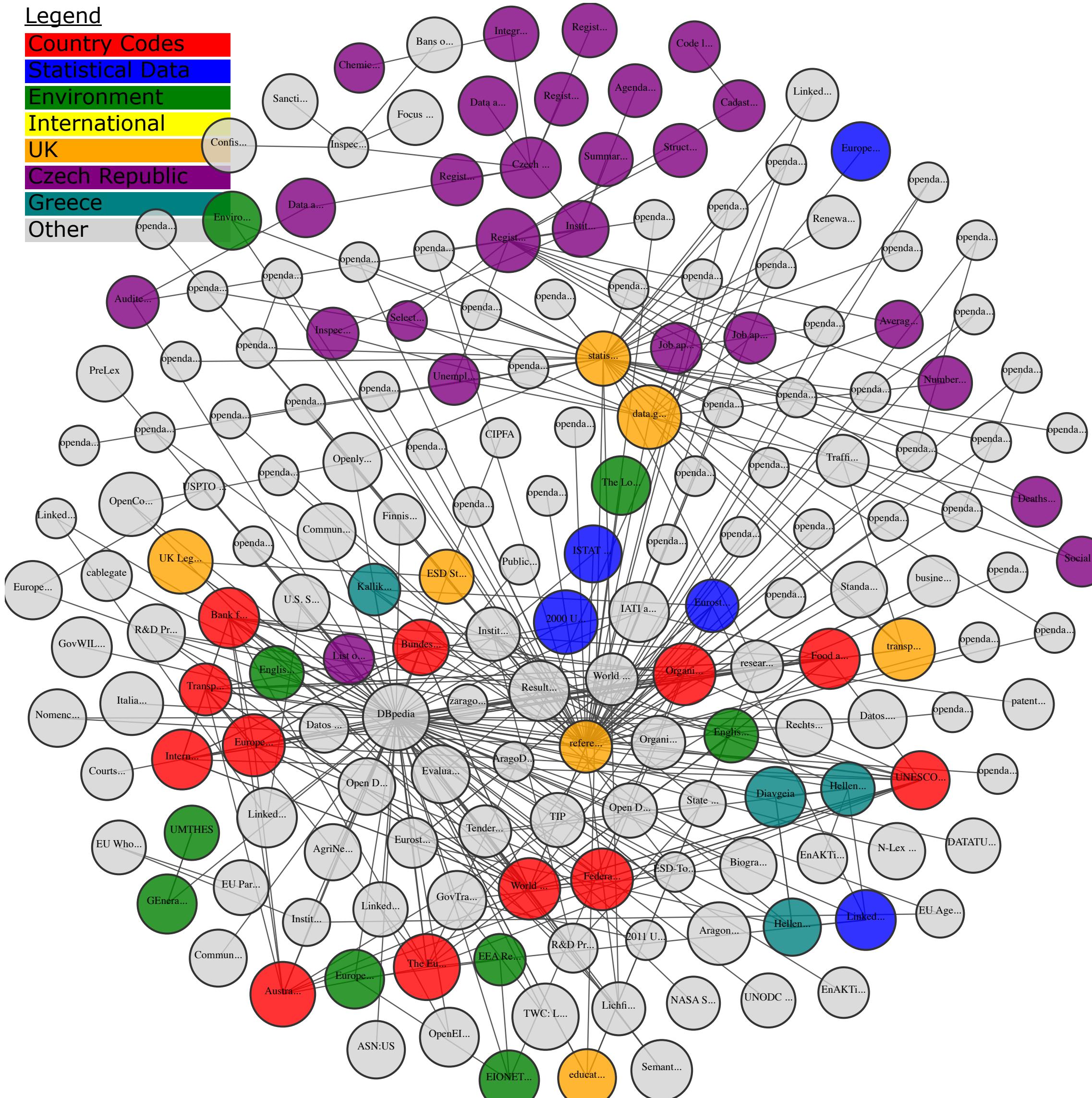


# March 2019

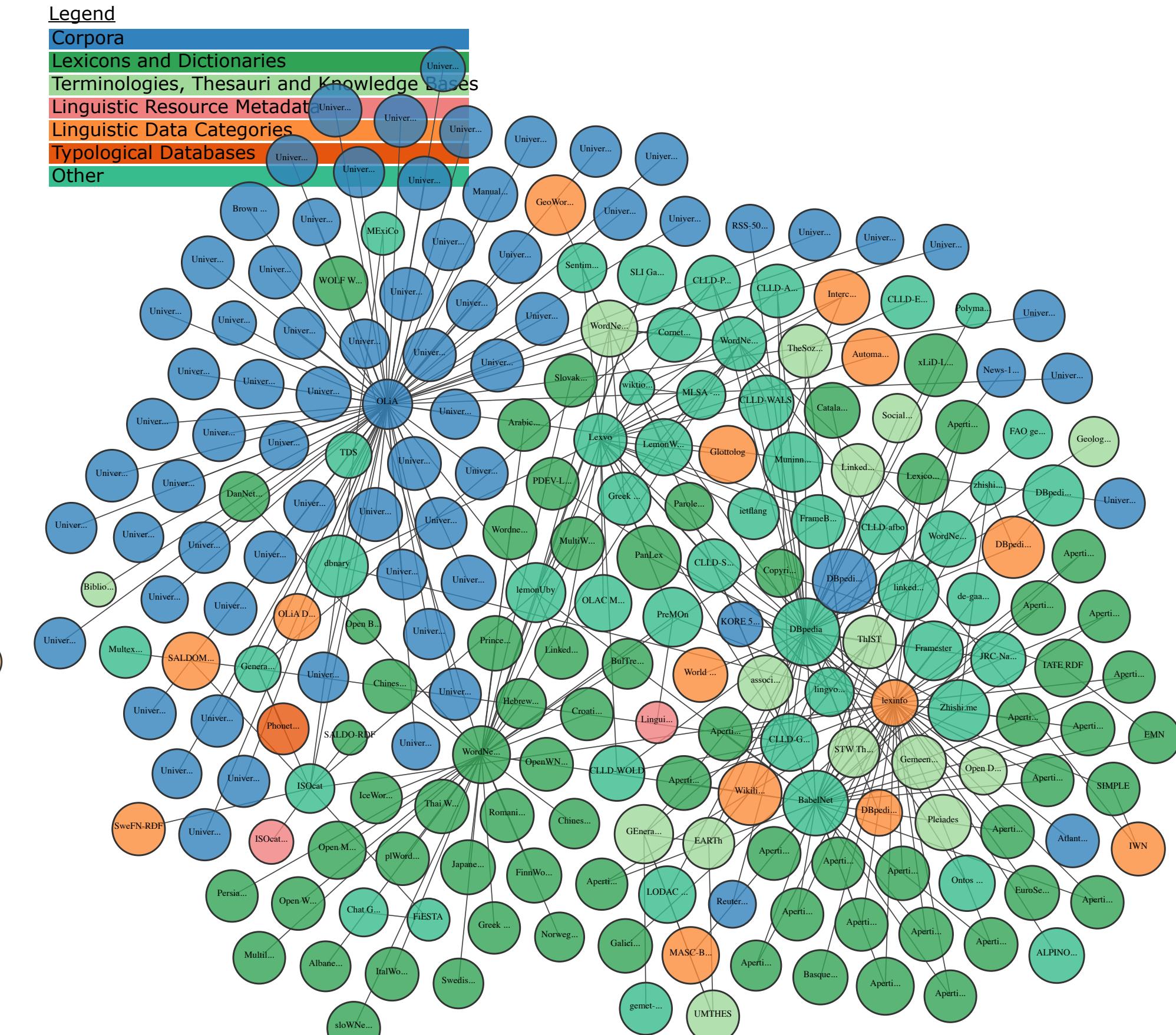
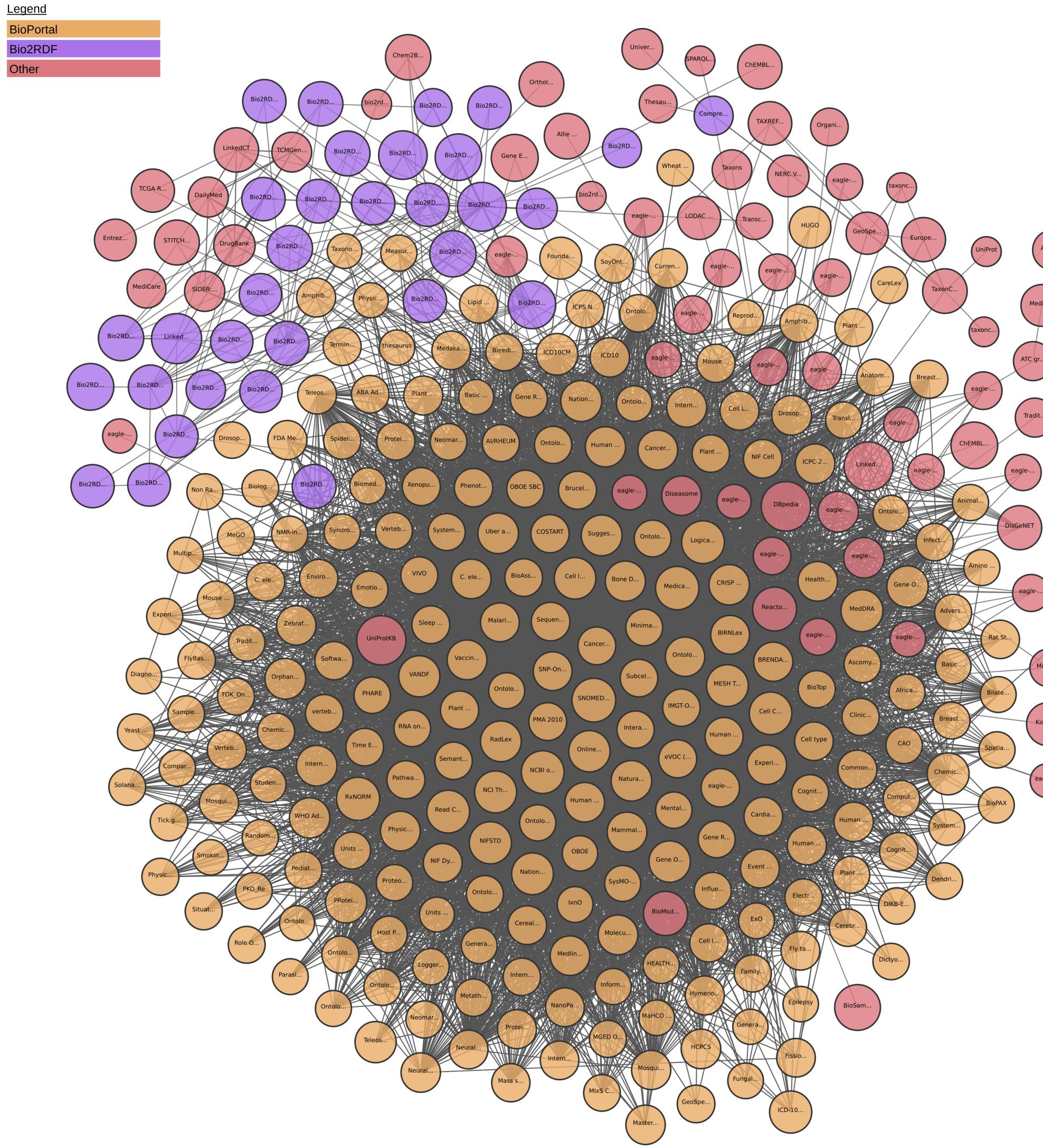
- FACTS:  
1,239 datasets with  
16,147 links!



# Government data



# Largest domain LODs



The Linguistic Linked Open Data Cloud from lod-cloud.net



# Where is the data

- Datasets
  - **UMBEL** - a lightweight reference structure of 20,000 subject concept classes and their relationships derived from OpenCyc, which can act as binding classes to external data; also has links to 1.5 million named entities from DBpedia and YAGO.
  - **FOAF** - a dataset describing persons, their properties and relationships
  - **VIAF (Virtual International Authority File)** - an aggregation of authority files (author names) from national libraries from around the world.
  - **Wikidata** - Wikidata is a free and open knowledge base, of more than 82,000,000 data items. It can be read and edited by both humans and machines.
  - **DBpedia** - a dataset containing extracted data from Wikipedia; it contains about 3.4 million concepts described by 1 billion triples, including abstracts in 11 different languages
  - **GeoNames** - provides RDF descriptions of more than 7,500,000 geographical features worldwide.

# LOD applications building blocks

## Using: Mashups

Mashups combine multiple datasets to create a new service, visualisation or information

## Using: Search

Linked data search engines allow search across the web of data. Conventional search may present information derived from linked data.

## Using: Productivity

Linked data facilitates data **integration** for business intelligence or research.

## Storing and publishing

Linked data can be published in simple flat files on a web server, in databases with a translation layer, or in specialised 'triple stores' built to store and share linked data.

## Querying: SPARQL

SPARQL Protocol and RDF Query Language provides a way to run structured queries over linked data datasets.

## Representing: Vocabularies

Vocabularies provide lists (and definitions) of common terms that can be used to describe the things and relationships in a dataset.

## Representing: Ontologies

Ontologies are vocabularies that record the logical relationships between their terms.

## Identifying: URLs

Using HTTP Uniform Resource Locators (URLs) means that (a) data can be looked up across the Internet; (b) decisions about 'namespaces' for data are managed through the Domain Name System (DNS).

## Interchanging: RDF

Resource Descriptor Framework (RDF) is a *model* for representing data as 'triples'. RDF can be serialised into a range of different file formats, including RDF-XML and text-based Turtle or N3 syntax.

## Transporting: HTTP (The World Wide Web)

Data is hosted on servers that can talk Hypertext Transfer Protocol (HTTP) to each other and to browsers in order to exchange data across the Internet.

## Licensing: open data

Open data is made available under licenses (or is placed in the public domain) so that others can use and build upon it, free of legal restrictions. Open standards for data files and interchange are used also.

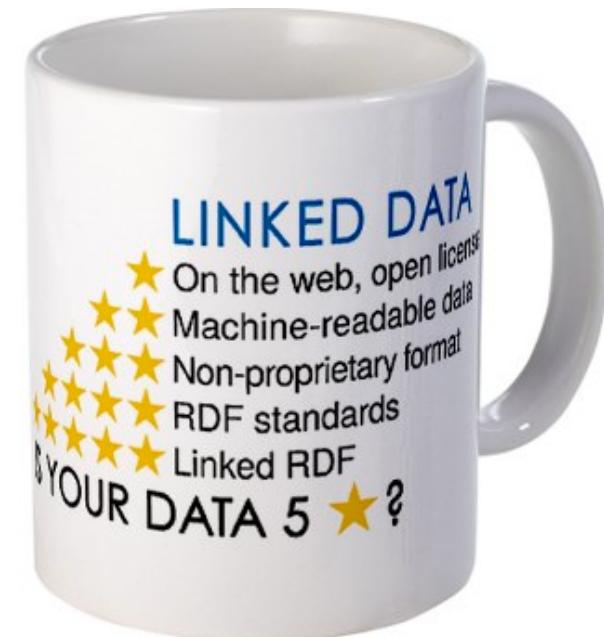
**F. Bauer, M. Kaltenböck**  
**Linked Open Data: The Essentials**  
**A quick Start Guide for Decision Makers**

Elements of the Linked Open Data Puzzle (revision 2) - 2nd May 2011. CC BY-SA-NC

Draft sketch by Tim Davies (@timdavies / tim@practicalparticipation.co.uk) for IKM Working Paper on Linked Open Data for Development. Comments welcome. Search 'linked open data puzzle' on <http://www.opendataimpacts.net> for latest version.

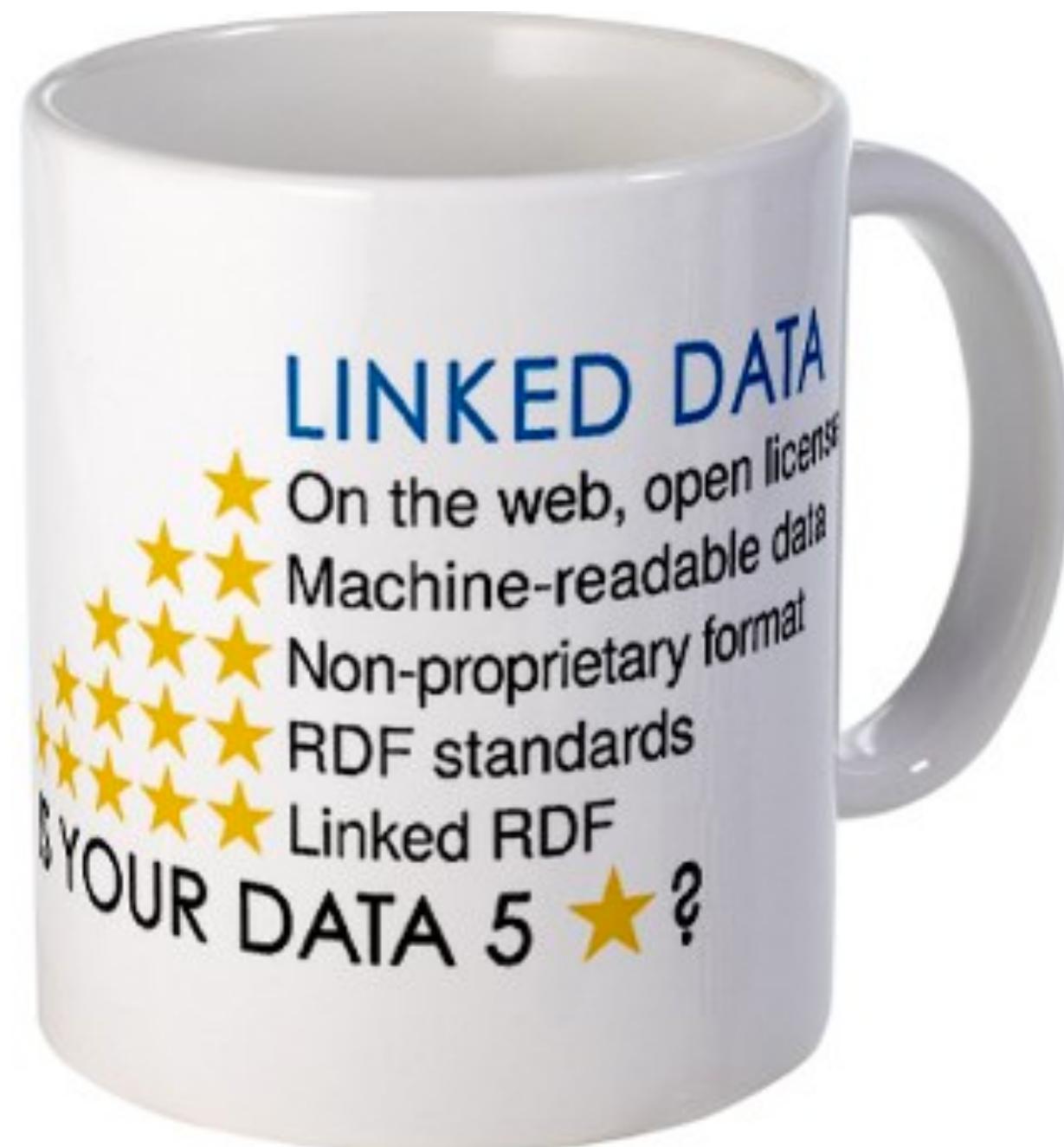
# How do we publish Linked Data

- Exposing Relational Databases or other similar formats into Linked Data
  - D2R
  - Triplify
  - R2O
  - NOR2O
  - Virtuoso
  - Ultrawrap
  - ...
- Using native RDF triplestores
  - Sesame
  - Jena
  - OwlIM
  - ...
- Incorporating it in the form of RDFa in CMSs like Drupal



# How do we consume Linked Data

- Linked Data browsers: To explore things and datasets and to navigate between them.
  - Tabulator Browser (MIT, USA), Marbles (FU Berlin, DE), OpenLink RDF Browser (OpenLink, UK), Zitgist RDF Browser (Zitgist, USA), Disco Hyperdata Browser (FU Berlin, DE), Fenfire (DERI, Ireland)
- Linked Data mashups: Sites that mash up (thus combine Linked data)
  - Revyu.com (KMI, UK), DBtune Slashfacet (Queen Mary, UK), DBPedia Mobile (FU Berlin, DE), Semantic Web Pipes (DERI, Ireland)
- Search engines: To search for Linked Data.
  - Falcons (IWS, China), Sindice (DERI, Ireland), MicroSearch (Yahoo, Spain), Watson (Open University, UK), SWSE (DERI, Ireland), Swoogle (UMBC, USA)



BETA This is a new service – your [feedback](#) will help us to improve it

# Search results

## Filter by

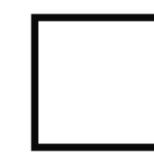
Publisher

Topic

Format

Open Government Licence  
(OGL) only

254 results found

Best match ▾

## [Organogram of Staff Roles & Salaries](#)

**Published by:** Serious Fraud Office

**Last updated:** 18 October 2016

Organogram (organisation chart) showing all staff roles. Names and salaries are also listed for the Senior Civil Servants. Organogram data is released by all central government departments and...

## [Organogram of Staff Roles & Salaries](#)

**Published by:** UK Supreme Court

**Last updated:** 18 October 2016

Organogram (organisation chart) showing all staff roles. Names and

# legislation.gov.uk

- Official government archive of the UK, managed by National archives.
- Best practice example of Open Government Data
  - Portal with access to all published UK legislation
  - Data from 1267 onwards, including recent changes in UK legislation
- Developed following W3C standards
- Persistent URIs for legislative data
  - identifier URIs, document URIs and representation URIs.

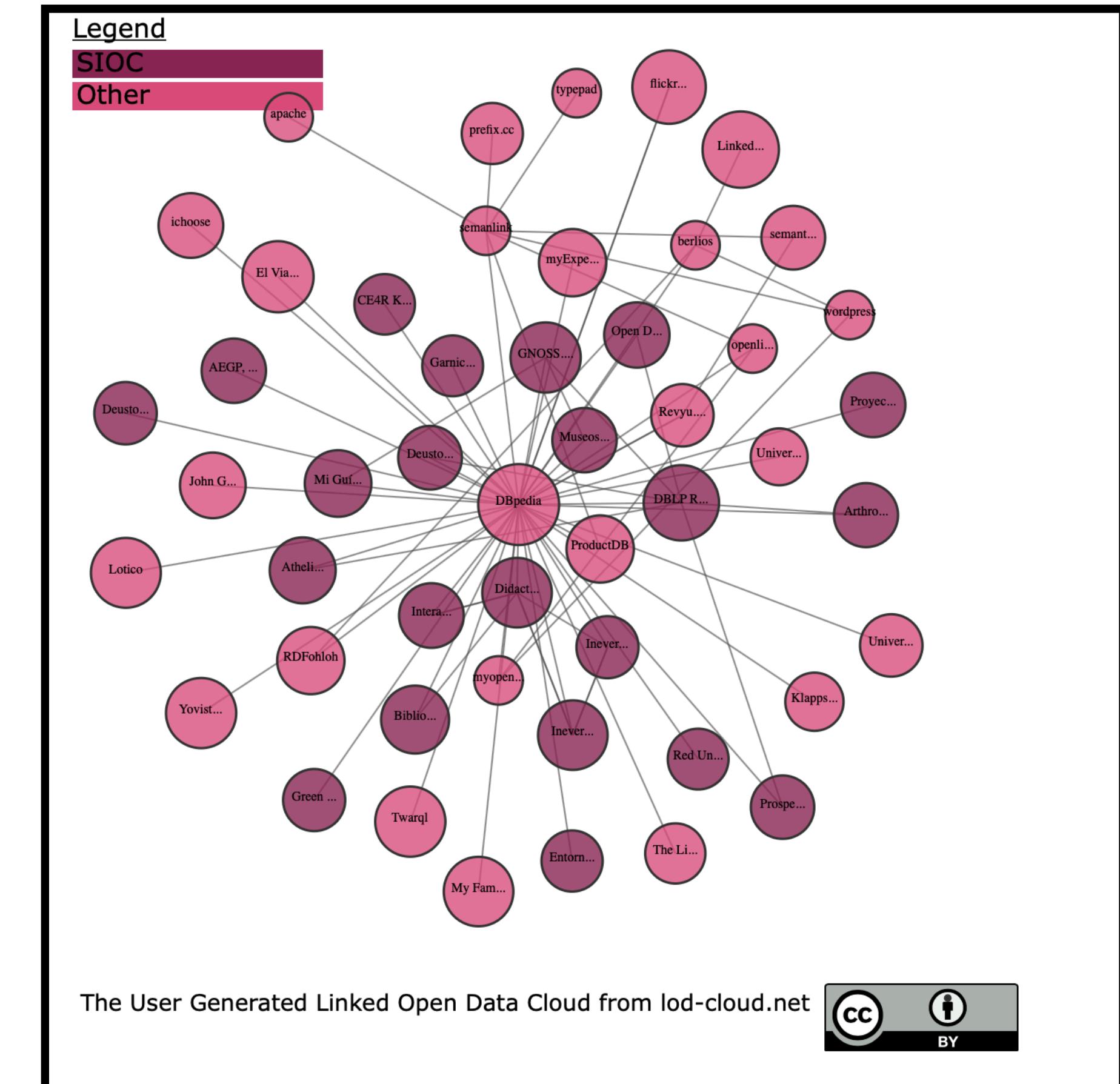
The screenshot shows the legislation.gov.uk homepage with a navigation bar at the top. Below the navigation, there's a search bar and a section titled 'URIs' which describes the Uniform Resource Identifier scheme used for legislation. It defines three levels of URLs: identifier URLs, document URLs, and representation URLs, with examples provided for each.

# Is it all good

- Well, there are downsides...
  - The quantity of published Linked Data increases day by day, but...
  - Some of the data available might be either irregularly updated, or already available in other formats and APIs often becomes an issue
    - This is not happening with all the datasets, but it needs to be taken under consideration.
  - Additionally, more data needs to be available to share, extent and re-use.
- Data should be urgently published as Linked Data on the Web with appropriate licenses and provenance information.
  - Risk of creating RDF silos.
- More applications and tools needed to exploit Linked Data.
- Existing open issues make the development of Linked Data based applications a challenge:
  - difficulties to integrate data in different formats and from multiple sources,
  - the discovery of data
  - or the usability of user interfaces.

# How to create Linked Data

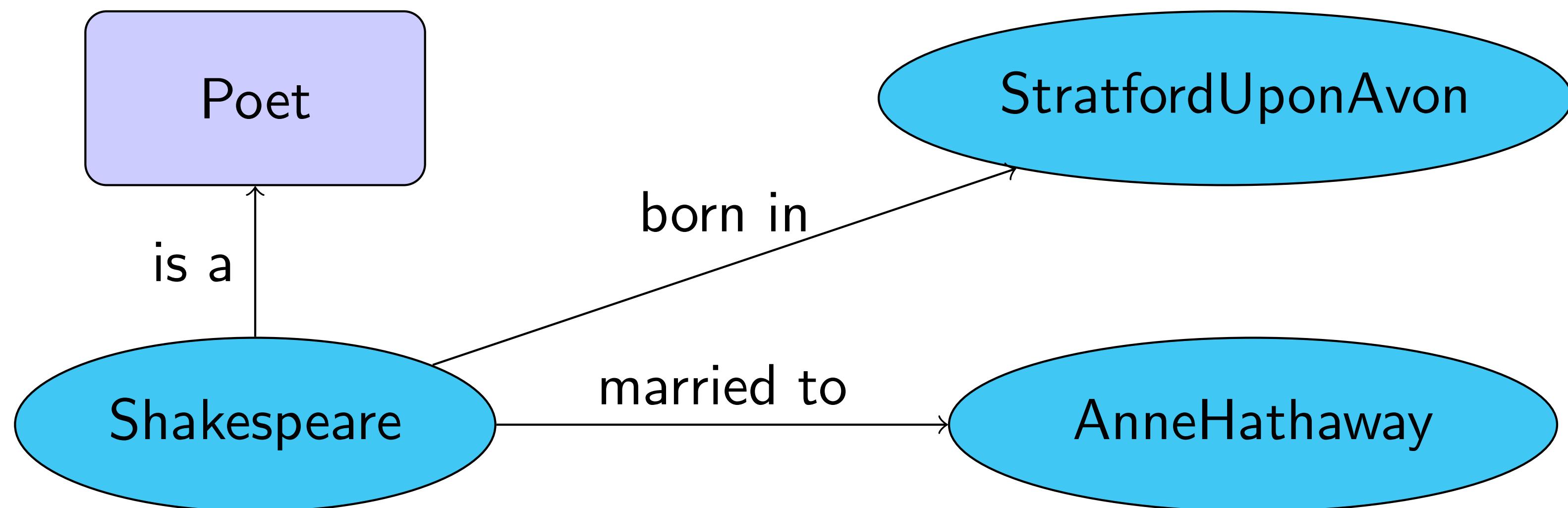
- 4-step creation process
  - Add semantics to the data
  - Model general world knowledge
  - Acquire new knowledge from inference
  - Query consistent information from different sources



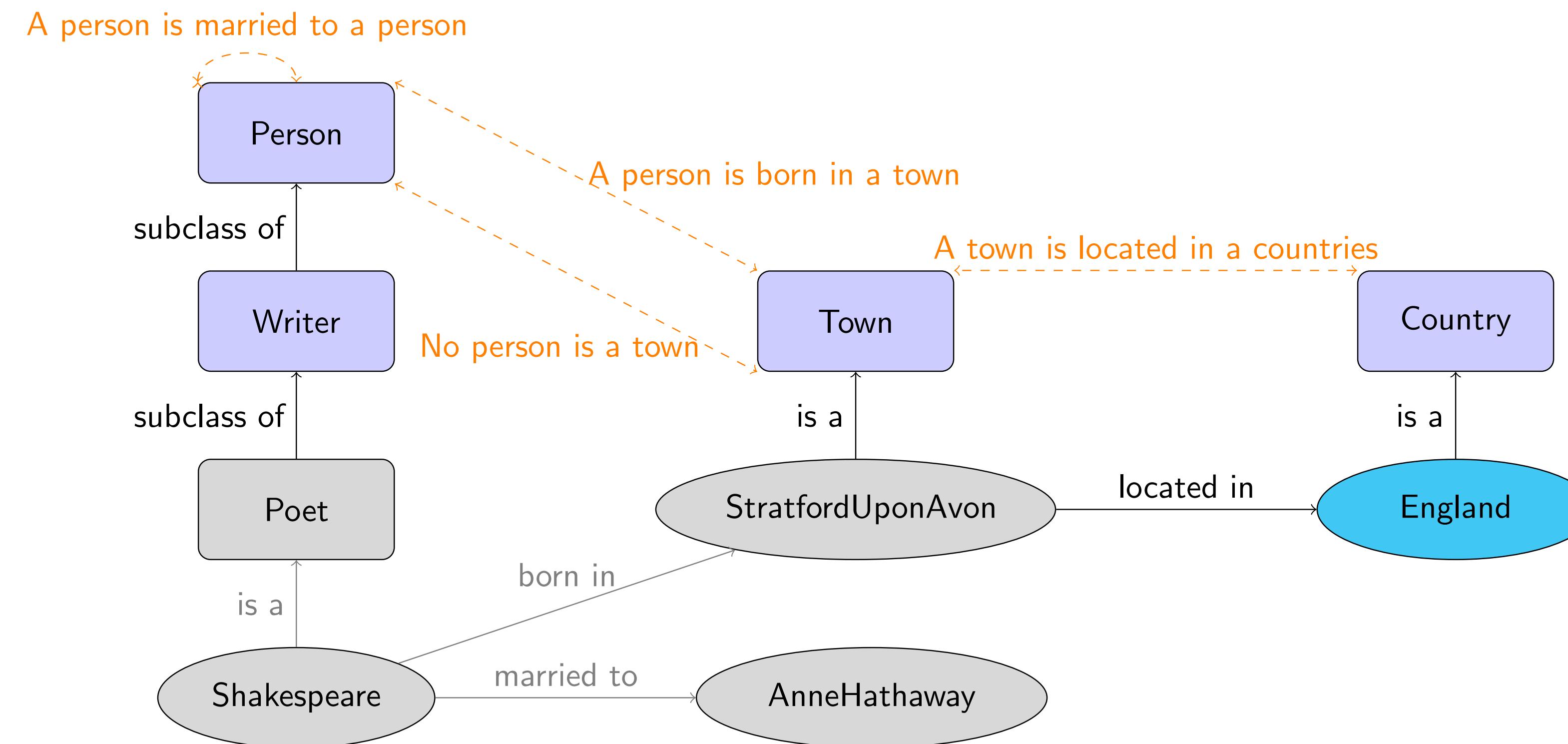
# Step 1: Add semantics to the data

William Shakespeare was an English poet and playwright. Shakespeare was born and brought up in Stratford-upon-Avon. At the age of 18, he married Anne Hathaway, with whom he had three children: Susanna, and twins Hamnet and Judith. ...

[Wikipedia]



# Step 2: Model general world knowledge



# Step 3: Acquire new knowledge from inference

- Is Shakespeare a writer?

- facts we know:
  - Shakespeare is a poet
  - Every poet is a writer

- facts we can infer or deduce →

- - Shakespeare is a writer

- Is Anne Hathaway a person?

- - Shakespeare is married to Anne Hathaway
  - A person is married to a person

- 

- - Anne Hathaway is a person

- Is Anne Hathaway a town?

- - Anne Hathaway is a person
  - No person is a town

- →

- - Anne Hathaway is NOT a town

# Step 4. Query information from different sources

- List all persons . . .
  - from Wikipedia.
  - from IMDB.
  - from Wikidata.
  - from British Library
- . . .
- List all people who are authors on a paper together with a co-author of Paul Erdoss.
  - from FOAF-profiles on the web
- List all stores who sell fair trade items from Ecuador.
- List all doctors who treat dermatitis.

# Recap

- Motivation for linked data
- Link data principles
- Link data adoption
- Designing of linked data