

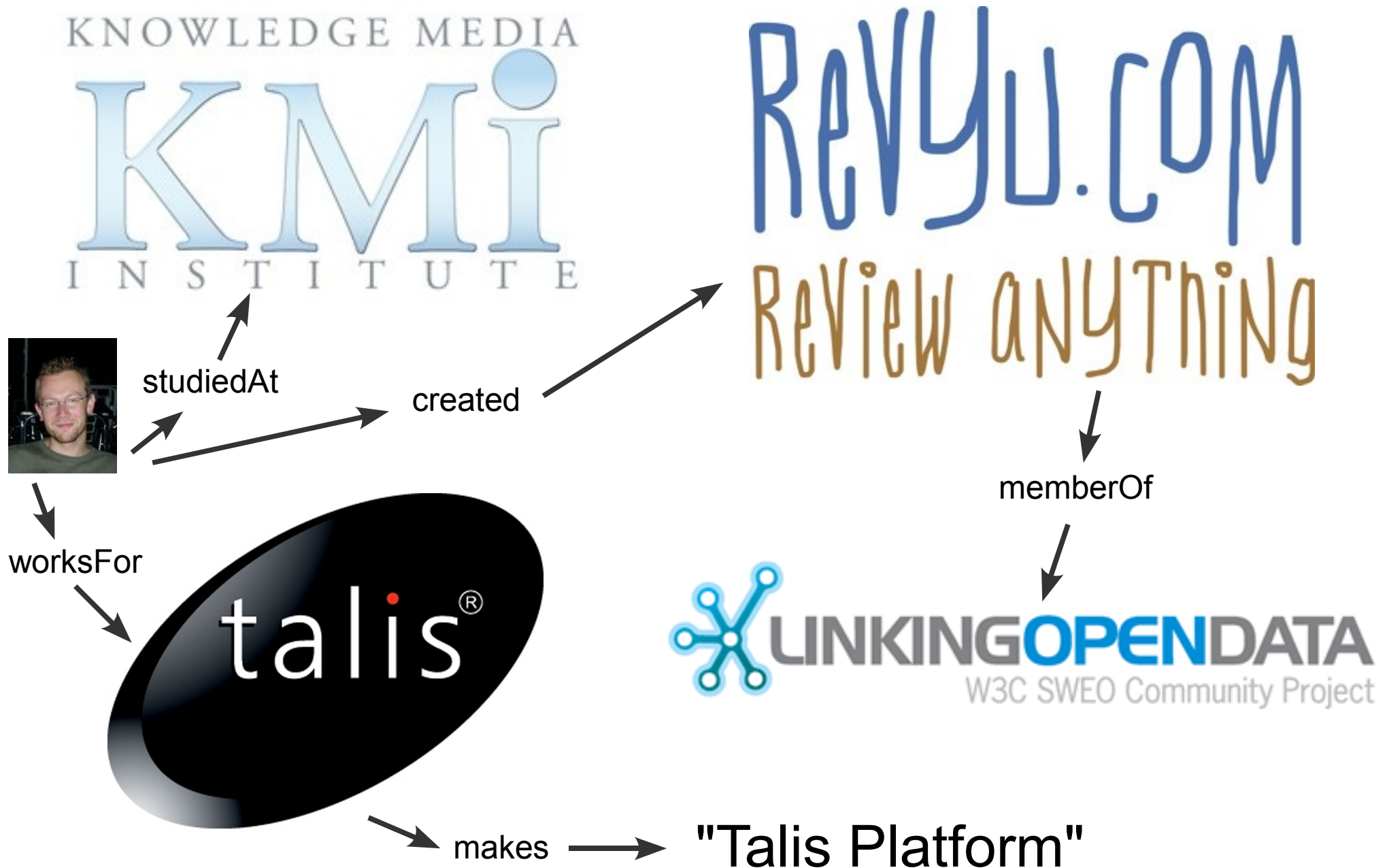
The Linking Open Data Project

Bootstrapping the Web of Data

Tom Heath
Talis Information Ltd, UK

CATCH Programme and E-Culture Project Meeting
on Metadata Interoperability
Amsterdam, 29 February 2008

My Background



- The Web of Documents and the Web of Data
 - From global filesystem to global database
- The Linking Open Data Project
 - Bootstrapping the Web of Data
- Getting Involved

The Web of Documents and the Web of Data

The Web of Documents

- Analogy
 - a global **filesystem**
- Designed for
 - **human** consumption
- Primary objects
 - **documents**
- Links between
 - **documents** (or sub-parts of)
- Degree of structure in objects
 - fairly **low**
- Semantics of content and links
 - **implicit**



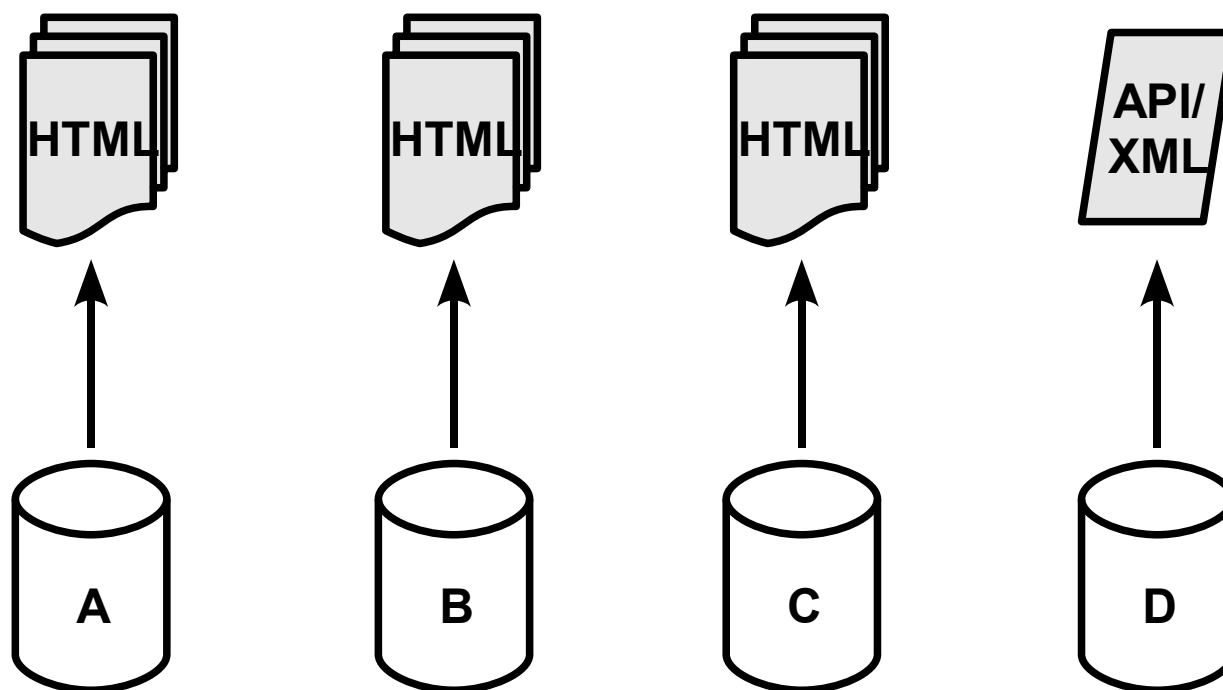
The Web of Documents: Issues

- **Simplicity**
 - Loosely structured data, untyped links, disconnected data
- **Integration**
 - Show me all the publications by EPSRC-funded PhD students
- **Querying**
 - Which papers have I written with people from European institutions outside the UK?

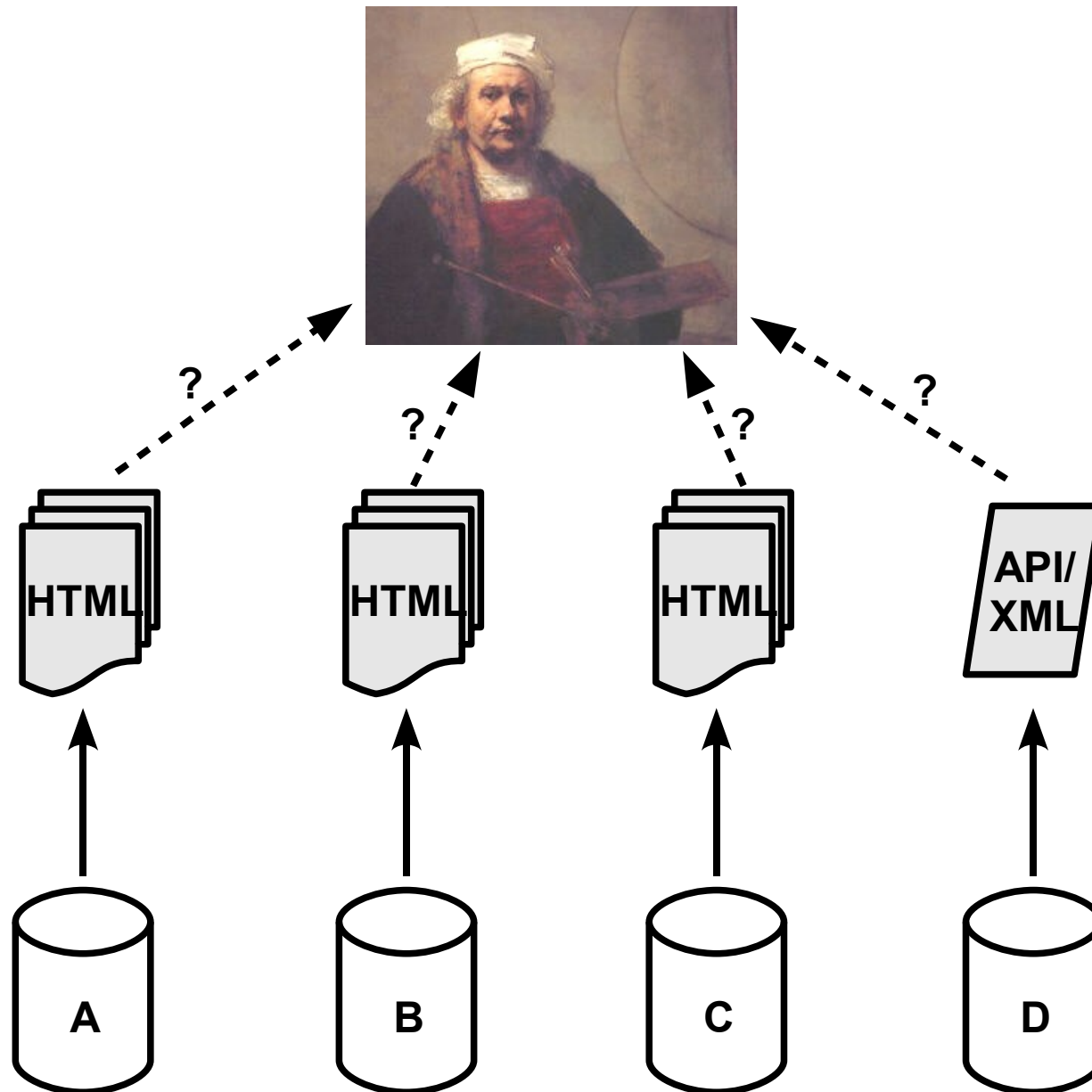
Data Silos on the Web



Data Silos on the Web



How do you identify Rembrandt on the Web of Documents?



Shared Identifiers support Data Interoperability

- Many common concepts or things need identifiers
- Reusing identifiers links data sets
- Linked data opens the doors of the silos and enables network effects

The Web of Linked Data

- Analogy
 - a global **database**
- Designed for
 - **machines** first, **humans** later
- Primary objects
 - **things** (or **descriptions** of things)
- Links between
 - **things**
- Degree of structure in (descriptions of) things
 - **high**
- Semantics of content and links
 - **explicit**



RDF: The Resource Description Framework

- Statements about things
- Triples: subject – predicate – object

<tom> <hasPet> <rover>

<rover> <type> <dog>

<rover> <colour> <brown>

The Linking Open Data Project

The Linking Open Data Project

- Problem

- it's getting boring playing with toy examples – we need real data to work with

- Solution

- take existing open data sets, convert them to RDF, publish them on the Web and link them together

The Linking Open Data Project

- Started February 2007 by Chris Bizer and Richard Cyganiak
- Supported by the W3C SWEO
- Current Participants
 - Universities
 - FU Berlin, MIT, KMi/The Open University, Universities of Pennsylvania, Leipzig, London, Hannover, Galway, Southampton, Karlsruhe...
 - Companies
 - OpenLink Software, Talis, Zitgist, Joanneum, BBC, Mondeca...
 - Outreach
 - Tim Berners-Lee, Ivan Herman (W3C), everyone...

Linked Data Principles

1. Use URIs to identify things

```
<http://tomheath.com/me>
```

2. Use HTTP URIs so people can look things up

```
GET /me HTTP/1.0
```

3. Provide useful data in RDF (preferably reusing ontologies)

```
<http://tomheath.com/me> rdf:type foaf:Person
```

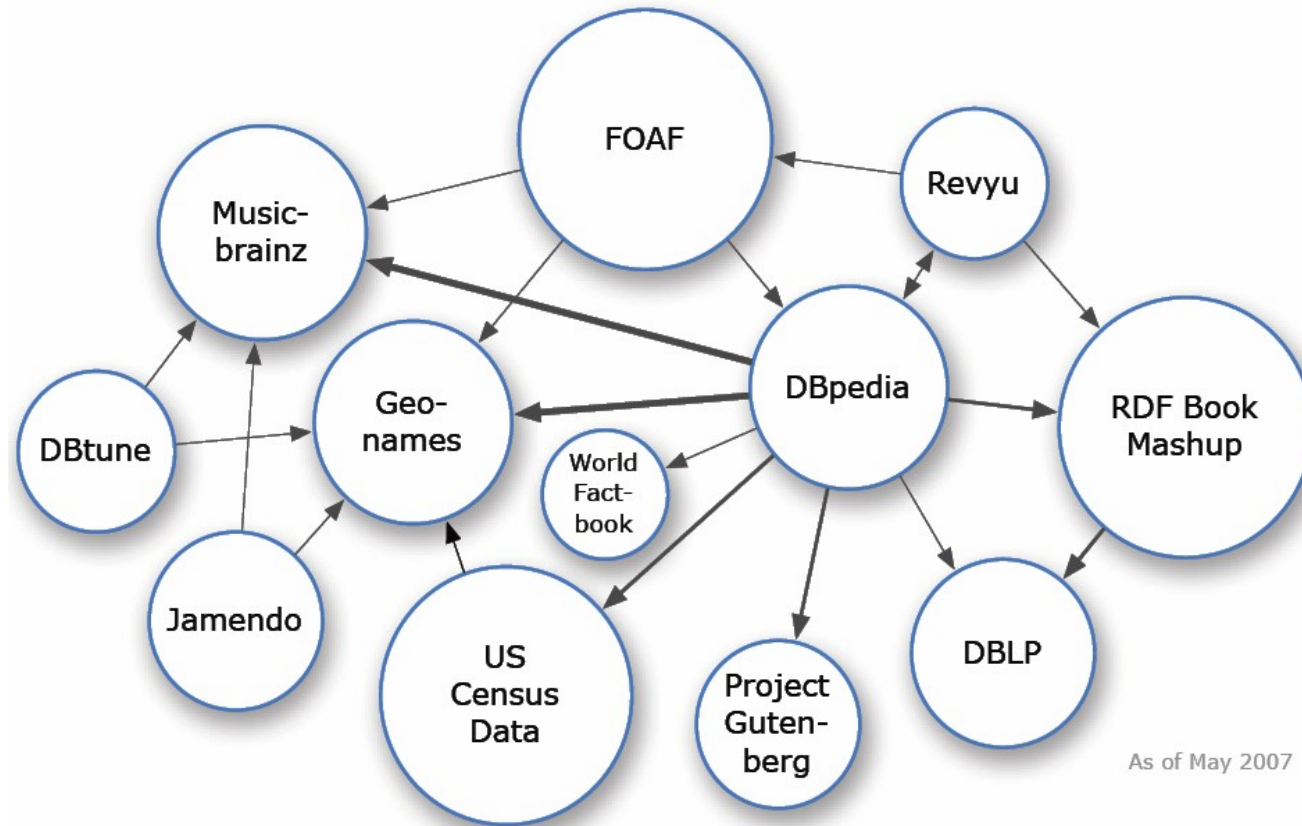
4. Use RDF to link to other things

```
<http://tomheath.com/me>
```

```
  eg:flewInto
```

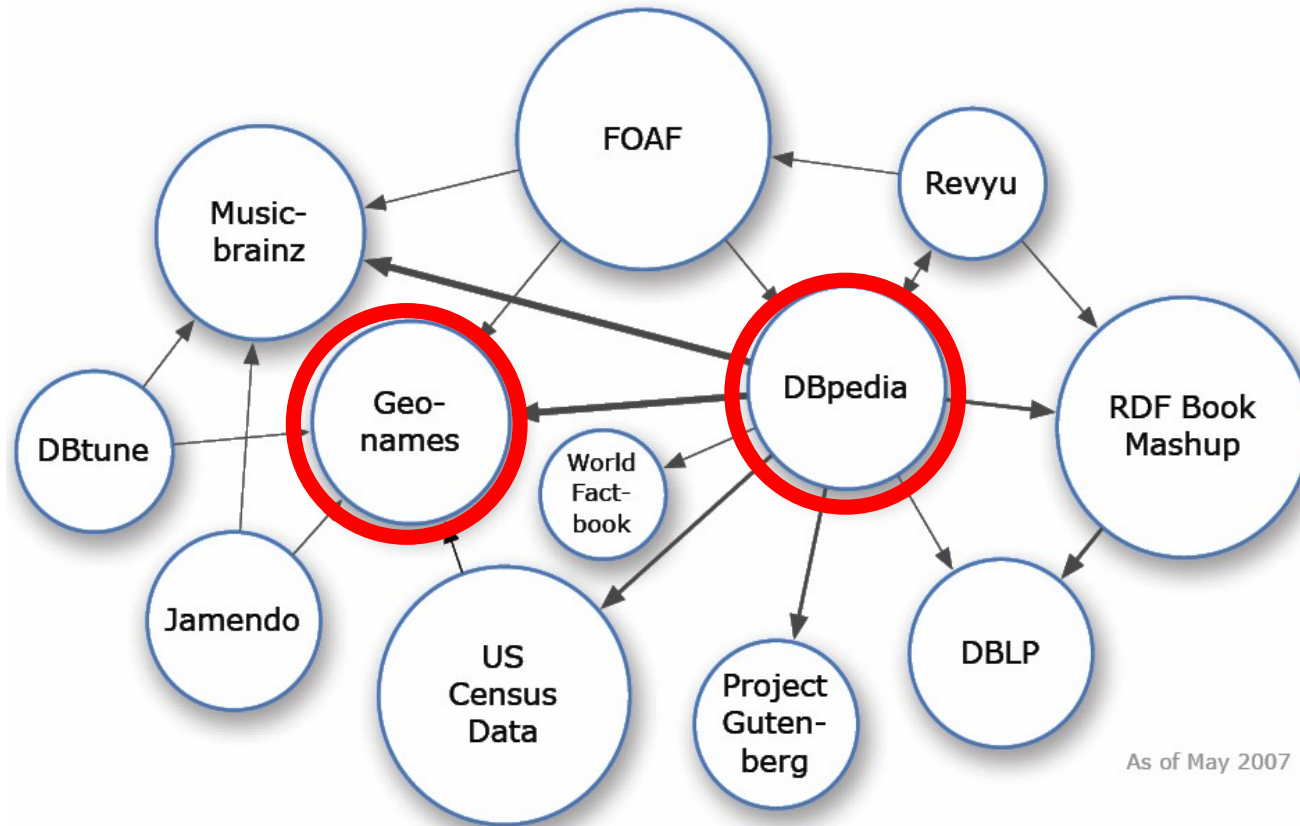
```
    <http://sws.geonames.org/6296680/>
```


The LOD "Cloud" - May 2007



Over 1 billion RDF triples served on the Web
Around 120,000 RDF links between data sources

The LOD "Cloud" - May 2007



Over 1 billion RDF triples served on the Web
Around 120,000 RDF links between data sources

Spotlight: DBpedia

- extract structured information from Wikipedia
- make this information available on the Web under an open license

Calgary	
	
Downtown Calgary.	
Government	
- Mayor	Dave Bronconnier (Past mayors)
- Governing body	Calgary City Council
- Manager	Owen A. Tobert
Area ^[1]	
- City	726.50 km² (280.5 sq mi)
- Metro	5,107.43 km² (1,972 sq mi)
Elevation	1,048 m (3,438.3 ft)
Population (2006) ^[1]	
- City	988,193
- Density	1,360.2/km² (3,522.9/sq mi)
- Metro	1,079,310
- Population rank	3rd
- Metro rank	5th

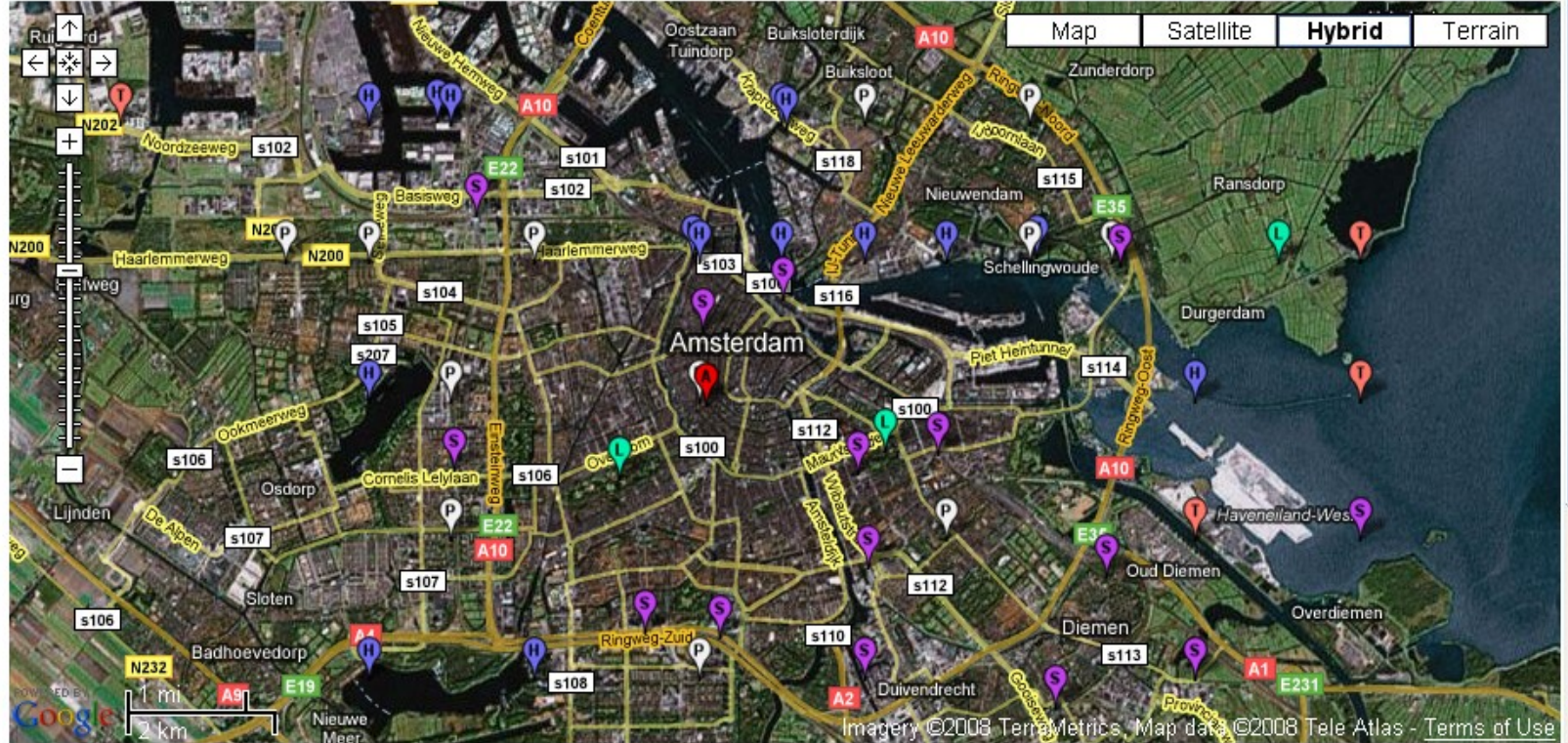
<http://en.wikipedia.org/wiki/Calgary>

```
<http://dbpedia.org/resource/Calgary>
  dbpedia:native_name "Calgary" ;
  dbpedia:altitude "1048" ;
  dbpedia:population_city "988193" ;
  dbpedia:population_metro "1079310" ;
  mayor_name
    dbpedia:Dave_Bronconnier ;
  governing_body
    dbpedia:Calgary_City_Council ;
  ...
```


Spotlight: Geonames

- Contains over eight million geographical names
 - 6.5 million unique features
 - 2.2 million populated places and 1.8 million alternate names
 - features categorized into one out of nine feature classes
 - further subcategorized into one out of 645 feature codes

Map center : N 52° 22' 13" E 4° 54' 3"

[ExploreOurPla.net](#) [google earth](#) [tagzania](#) [mapquest](#)[GeoNames](#) [Wikipedia](#)

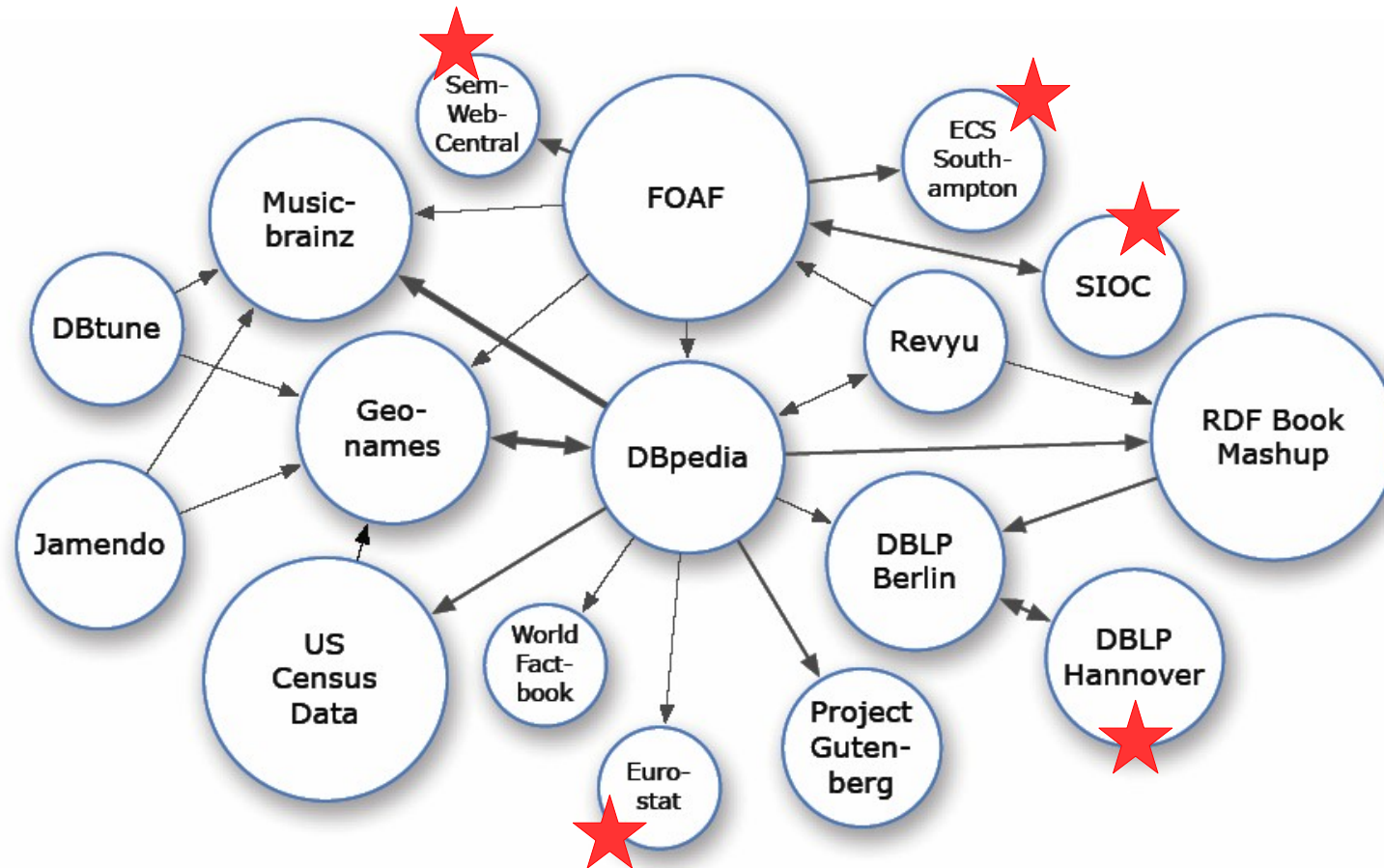
features

- ☒ city, village, ...
- ☒ mountain, hill, rock, ...
- ☒ stream, lake, ...
- ☒ country, state, region, ...
- ☒ parks, area, ...
- ☒ road, railroad
- ☒ spot, building, farm
- ☒ forest, heath, ...
- ☒ undersea

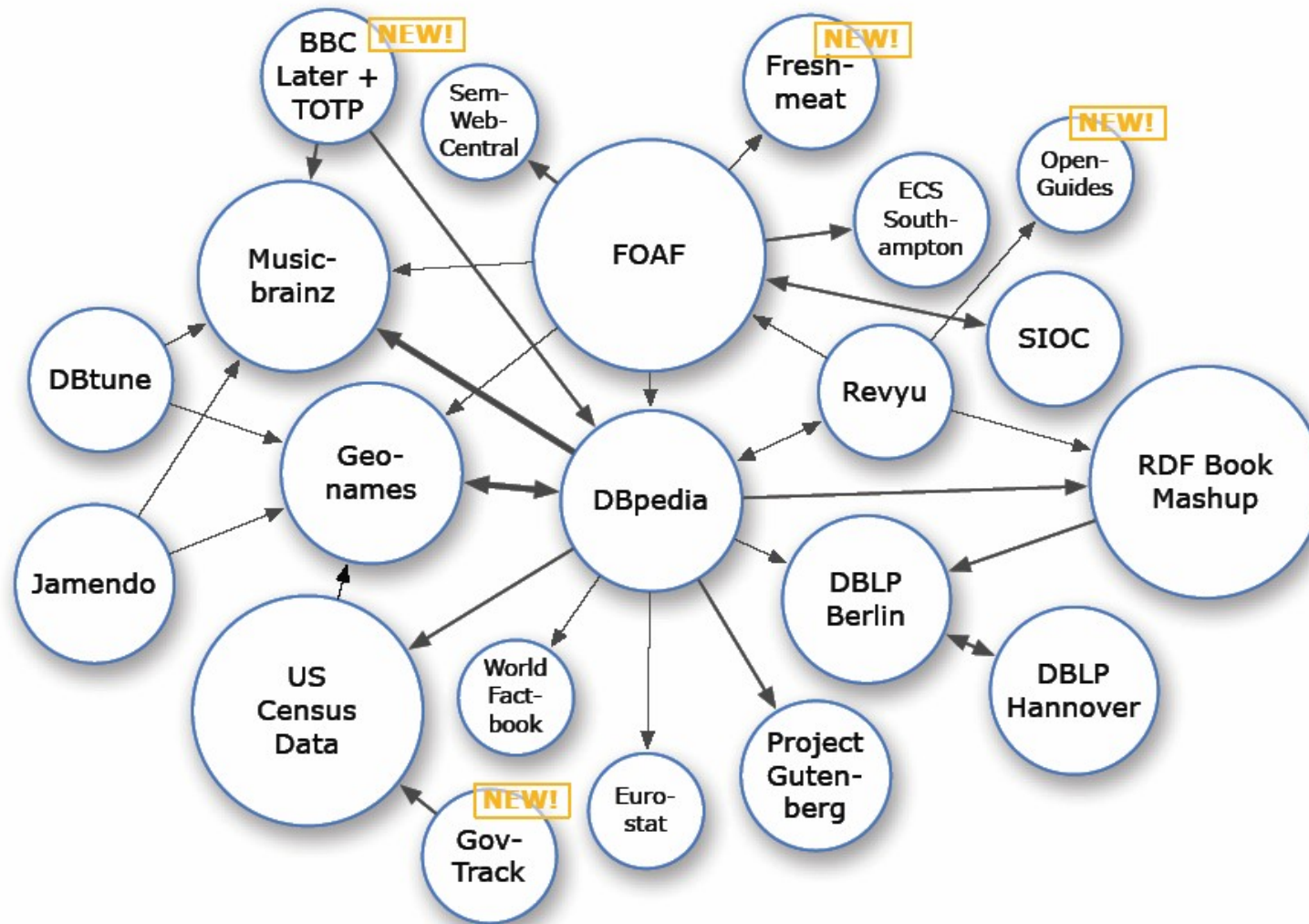
only 50 objects displayed, zoom in or deselect some features

	Name	country	feature	km to center
1	Amsterdam Centraal Station	Netherlands	railroad station	0.96 km
2	Anne Frank House	Netherlands	museum	1.25 km
3	Jordaan	Netherlands	section of populated place	1.26 km
4	Gemeente Amsterdam	Netherlands	second-order administrative division	1.26 km
5	Het IJ	Netherlands	navigation channel	1.44 km
6	Onze Lieve Vrouwe Gasthuis Locatie Oosterpark	Netherlands	clinic	1.68 km
7	Oosterpark	Netherlands	park	1.73 km
8	Dijksgracht	Netherlands	docking basin	1.79 km
9	Westerkanaal	Netherlands	canal	1.87 km

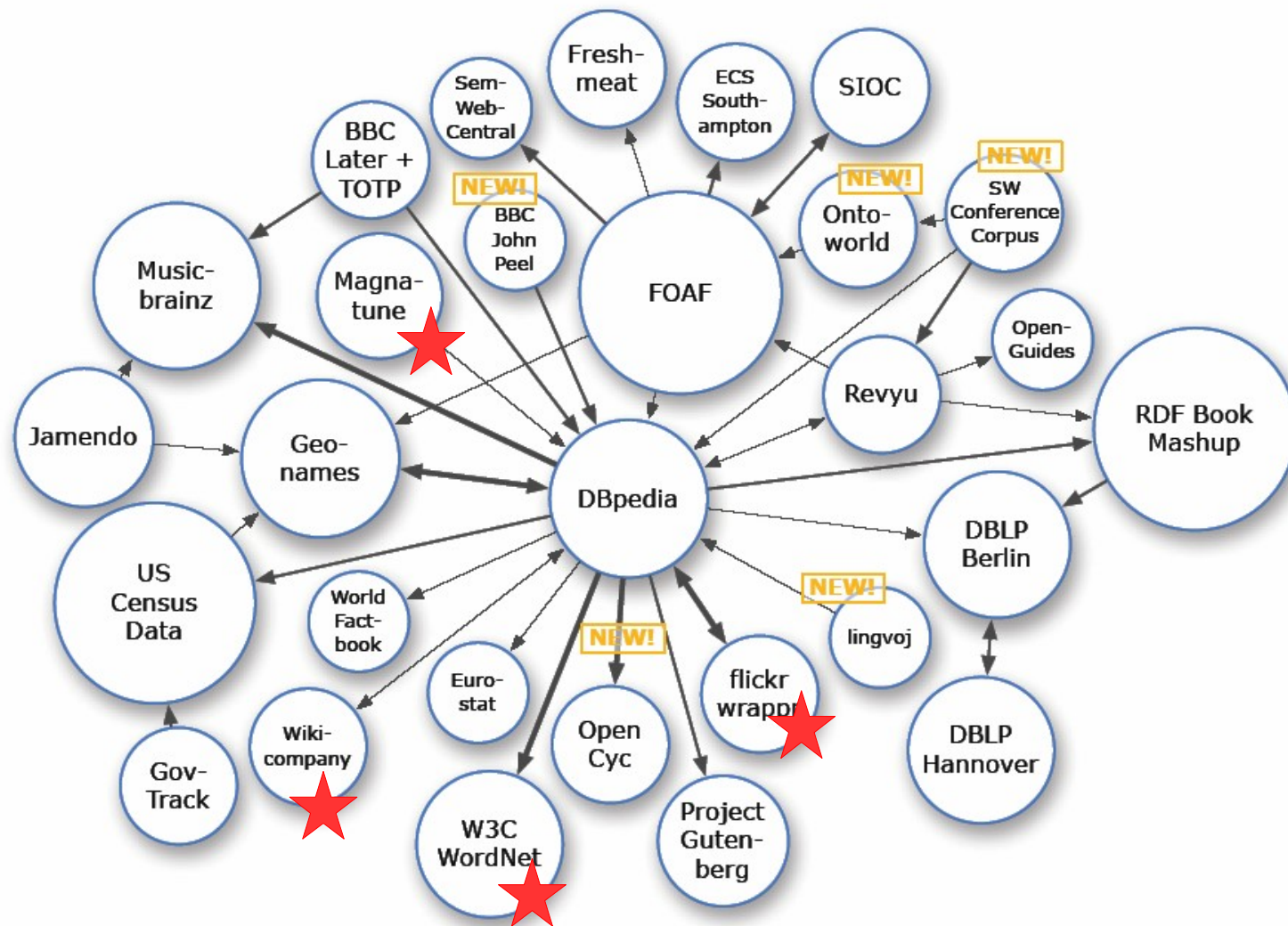
The LOD "Cloud" - July 2007



The LOD "Cloud" - August 2007

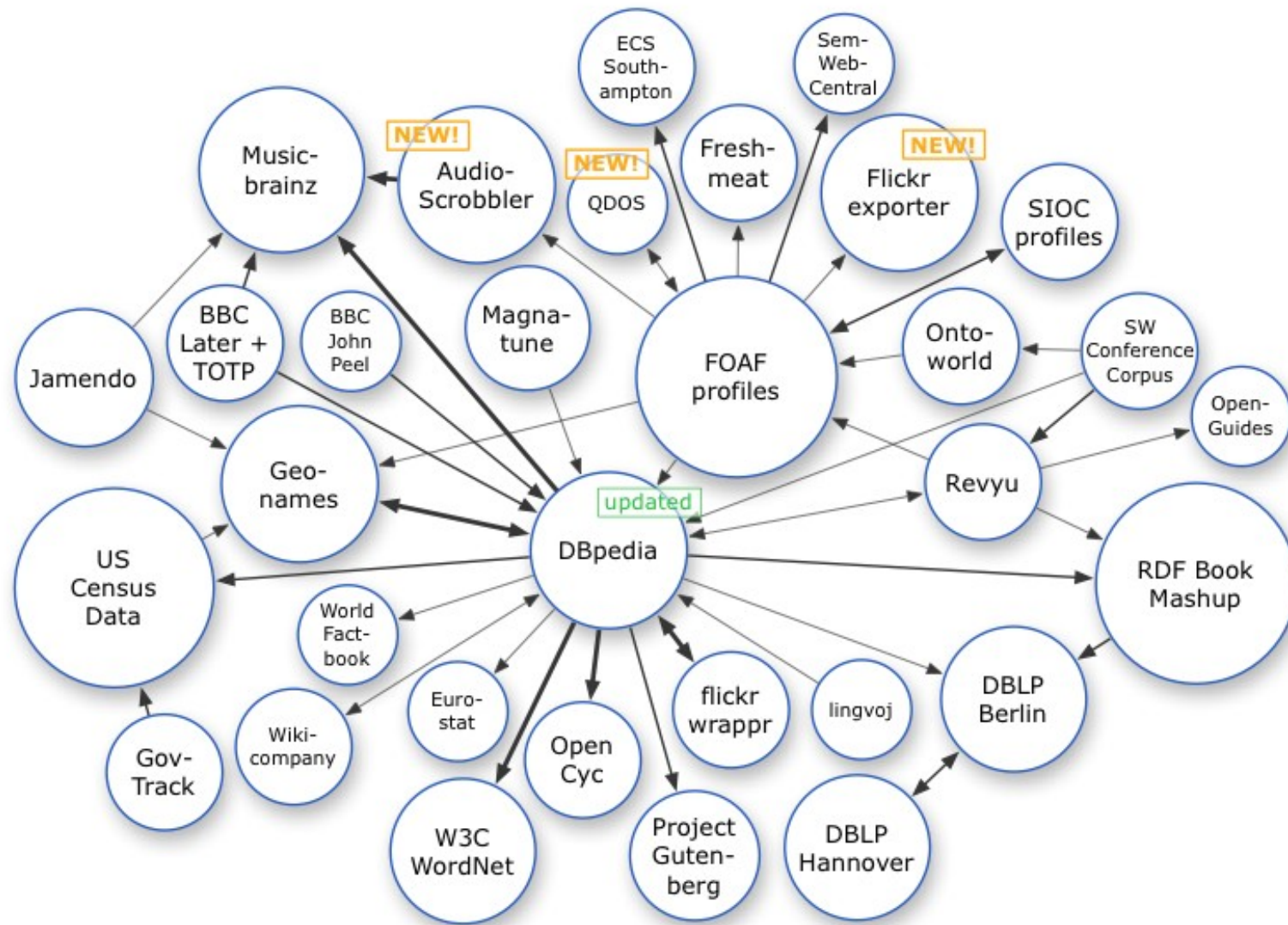


The LOD "Cloud" - Nov 2007



Over 2 billion RDF triples served on the Web
Around 3 million RDF links between data sources

The LOD "Cloud" – Feb 2008




Linked Data Applications

Linked Data Browsers

Disco - Hyperdata Browser ([About](#))

Berlin

URI:

Property	Value	Sources
population	3398888	G2
type	http://dbpedia.org/Category/City	G2
comment	Berlin is the capital city and one of the sixteen Federal States of Germany. It is the country's largest city in area and population, and the second most populous city in the European Union.	G2
comment	Berlin ist die deutsche Bundeshauptstadt und als Stadtstaat ein eigenständiges Land der Bundesrepublik Deutschland. Berlin ist die bevölkerungsreichste und flächengrößte Stadt Deutschlands und nach Einwohnern die zweitgrößte Stadt der EU.	G2
label	Berlin	G2
sameAs	http://sws.geonames.org/2950159/	G2
subject	http://dbpedia.org/resource/Category/Berlin	G2
subject	http://dbpedia.org/resource/Category/Capitals_in_Europe	G2
subject	http://dbpedia.org/resource/Category/Cities_in_Germany	G2
subject	http://dbpedia.org/resource/Category/German_state_capitals	G2
subject	http://dbpedia.org/resource/Category/Host_cities_of_the_Summer_Olympic_Games	G2
subject	http://dbpedia.org/resource/Category/States_of_Germany	G2
sourceURL	Berlin	G1
depiction		G2
page	http://en.wikipedia.org/wiki/Berlin	G2
is birthplace of	Adolf von Baeyer	G2
is birthplace of	http://dbpedia.org/resource/Person/Albert_Speer_%28the_younger%29	G2

Linked Data Mashups – Revyu

Broken Flowers

Links

Homepage: <http://www.brokenflowersmovie.com/>

See Also: http://en.wikipedia.org/wiki/Broken_flowers

Tags

[bill-murray](#) [film](#) [jessica-lange](#) [jim-jarmusch](#) [julie-delpy](#) [movie](#) [sharon-stone](#)

Reviews (1)

★★★★★ [by tom on 30 Jan 2007](#)

Broken Flowers provides a fantastic vehicle for a classic deadpan Bill Murray performance. The film centers around his character Don, who one day receives a letter from an ex-girlfriend, telling him he has a teenage son. The letter is unsigned, so (with encouragement from his neighbour) he sets off round the country, visiting each the exes who could be the mother of his son. Predictably they're all different in personality and life situation, giving plenty of raw material for awkward silences and dubious encounters. This is great viewing for any Bill Murray fans, or anyone who likes their humour intelligent and a little bit quirky. The soundtrack is also excellent, and deserves a separate review.

What do you think of **Broken Flowers**? [Write Your Own Review...](#)

Broken Flowers



directed by [Jim Jarmusch](#)

Linked Data Mashups – Revyu

Broken Flowers



directed by Jim Jarmusch

RDF Metadata About
Broken Flowers

Linked Data Mashups – Revyu

Reviews by *tom* (81)

The Acorn Inn, Evershot, Dorset

★★★★☆ on 14 Feb 2007

The Acorn Inn isn't your typical village pub. It sits on the main street of the very pretty and very chocolate box village of Evershot, and features in the story Tess of the d'Urbervilles by Thomas Hardy. As well as the bar area at the back, the front is now a fairly high-end restaurant, so the place seems to want to be all things to all people. Generally it manages it pretty successfully.

We had a pub meal and a few beers in the bar. The food was good, tasty pub food (cod and chips, ham egg and chips), although there wasn't much choice on the bar meals menu itself (three main dishes plus sandwiches and lighter stuff). We had the option of ordering off the restaurant a la carte menu aswell, but weren't planning to go that upmarket. We did however order puddings off the restaurant menu (chocolate brownie, creme brulee). These were fantastic, presented restaurant style, almost as expensive as our main courses, but well worth it.

There were three different real ales on tap, with good variety of different types (two bitters and a stout), but the thing that impressed me most was the service. Both the waitress who served us and the barman were incredibly friendly and professional, but in a nicely laid back way. The one thing I'd change about the service was that we were

[Web Feed of **tom**'s Latest Reviews](#)

 WEB FEED

About **tom** (Tom Heath)



[tom's Home Page](#)

[RDF Metadata About **tom**](#)

 META 

Linked Data Mashups – Revyu

Web Feed of **tom's**
Latest Reviews



About **tom** (Tom
Heath)



tom's Home Page

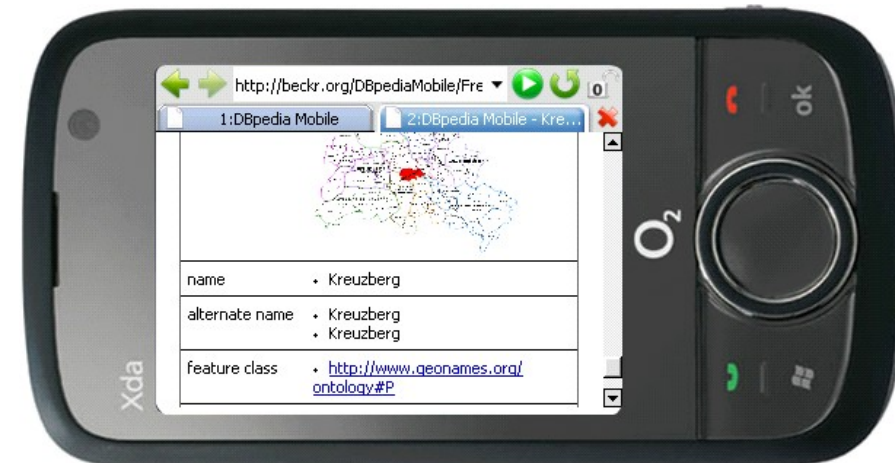
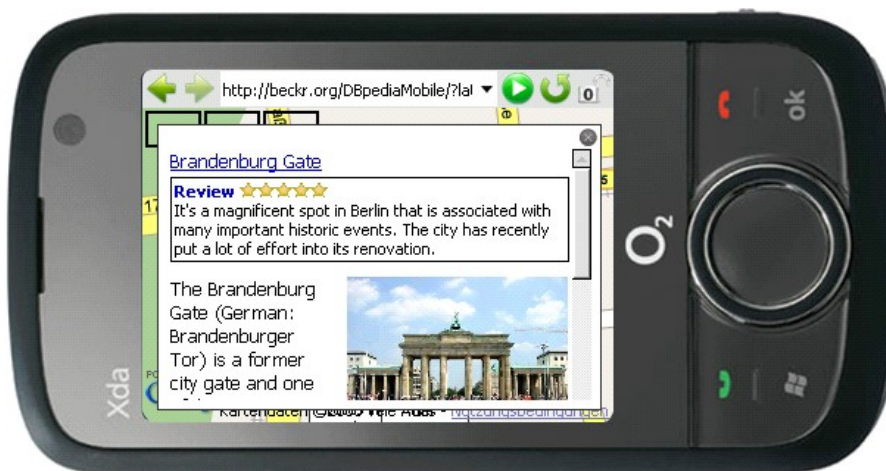
RDF Metadata About
tom



■ Add to del.icio.us

DBpedia Mobile

- Geospatial entry point into the Web of Data
- Uses DBpedia, Revyu and Flickr
- Credits: Christian Becker and Christian Bizer, FU Berlin



Outlook

Queries of the Future

- Whereabouts near my home can I see buildings by architects who were influenced by the Bauhaus?
 - ...on a Monday?
 - ...and with a student discount?



Queries of the Future

- Which European city has the greatest concentration of works by Caravaggio?
 - ...and has direct flights from my home town?
 - ...with an airline that is rated good or excellent?
 - ...by me? ...by my friends?



Getting Involved

Getting Involved

- Which data sets are you responsible for?
- How might these connect to existing "hubs" in the Web of Data?
- Which new "hubs" might you be able to create?
- Get more information via <http://linkeddata.org/>
- Add your name to the LOD wiki page
- Join the LOD mailing list and say "Hi"
- Link some data!

Thankyou – Any Questions?

- More info: <http://linkeddata.org/>
- My URI: <http://kmi.open.ac.uk/people/tom>
- Talis Platform: <http://www.talis.com/platform>
- Slides:
- <http://linkeddata.org/slides/2008-02-amsterdam-catch.pdf>