Linking Distributed Data across the Web

Dr Tom Heath

Researcher, Platform Division

Talis Information Ltd

tom.heath@talis.com
http://tomheath.com/

Overview

- Background

- From a Web of Documents to a Web of Data

- The Linking Open Data Project

- Applications of Linked Data

- Ongoing Challenges

About Talis

- Birmingham-based software company

- Past: established nearly 40 years ago to deliver large collaborative metadata systems

- Present: very stable with over 100 customers - public authorities and universities

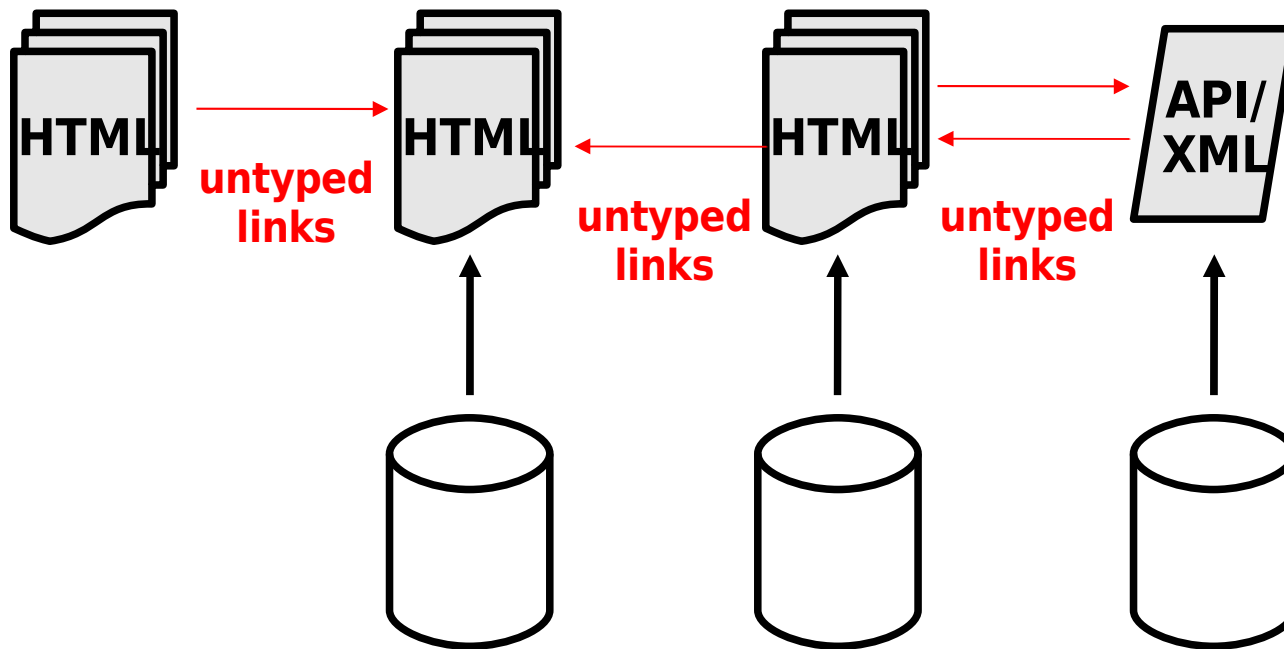- Future: global data platform supporting ours and others' applications

Web of Documents to 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 Linked Documents
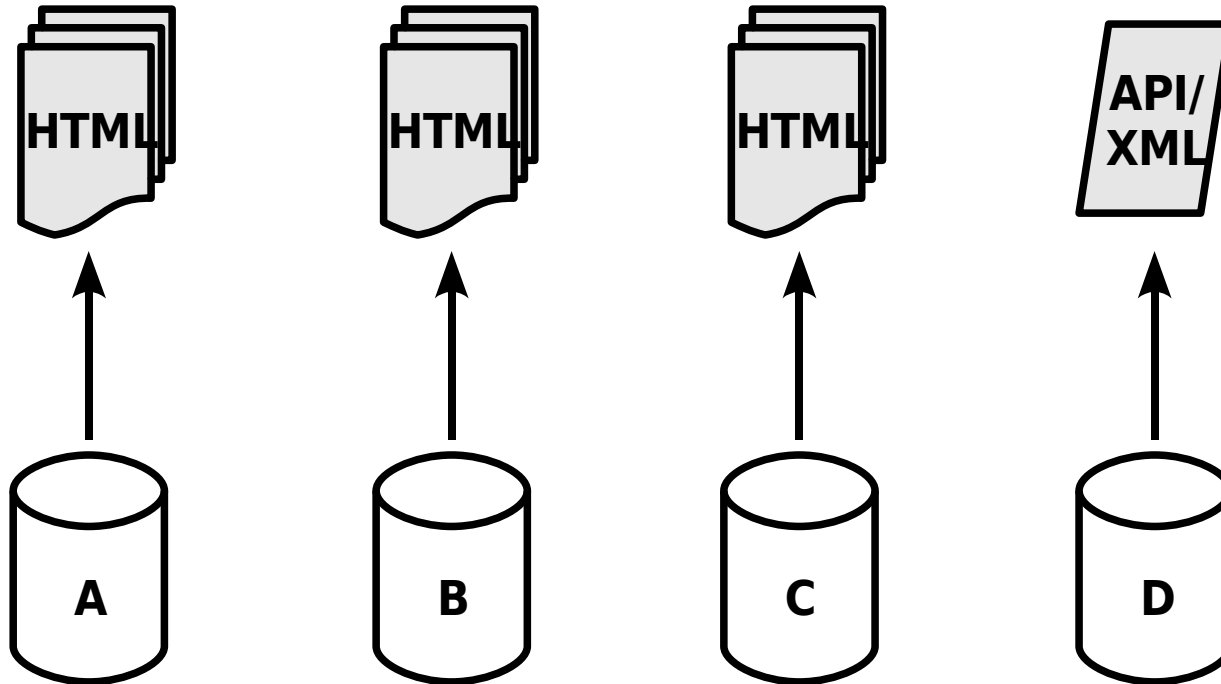
The Web of Documents: Issues

- Simplicity
  - Loosely structured data, untyped links, disconnected data

- Integration
  - Show me all the publications by publicly-funded PhD students

- Querying
  - Which papers have I written with people from European institutions outside the UK?

Data Silos on the Web

talis®

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**

The Web of Linked Data

Don't just link the *documents*, link the *things*

| Thing | | Thing | | Thing | | Thing | | Thing |

typed links   **typed links**   **typed links**   **typed links**
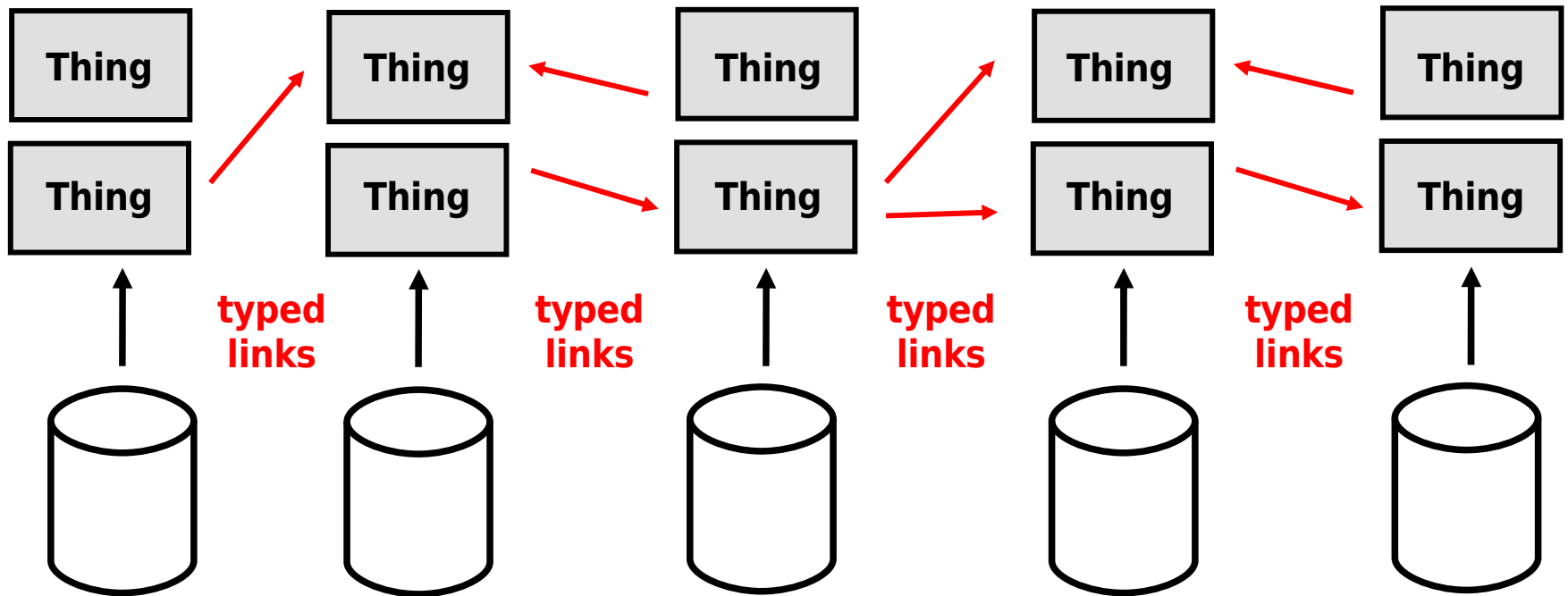
Linked Data is…

- …a way of publishing data on the Web that:

  – encourages reuse
  – reduces redundancy
  – maximises its (real and potential) inter-connectedness
  – enables network effects to add value to data

Linked Data Technology Stack

- URIs
- HTTP
- RDF
- (RDFS/OWL)

URIs – Not Just for Web Pages

- "A Uniform Resource Identifier (URI) provides a simple and extensible means for identifying a resource." -- RFC 3986

- Many different schemes: *http://, ftp://, tel:, urn:, mailto:*

- Some URIs for "real world" things:
  - http://tomheath.com/id/me
  - http://dbpedia.org/resource/Talis_Group
  - http://sws.geonames.org/4671654/

HTTP

- Data access mechanism

- Using *http://* URIs to identify things allows people to look these things up

RDF: Resource Description Framework

- Data format for describing things and their interrelations

The RDF Data Model

- *Triples*

subject → predicate → object

Tom → worksFor → Talis

Talis → basedIn → Birmingham

`<uri>  →  <uri>  →  <uri> or "literal"`

"Talis is Based Near Birmingham"

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

<http://xmlns.com/foaf/0.1/Person#based_near>
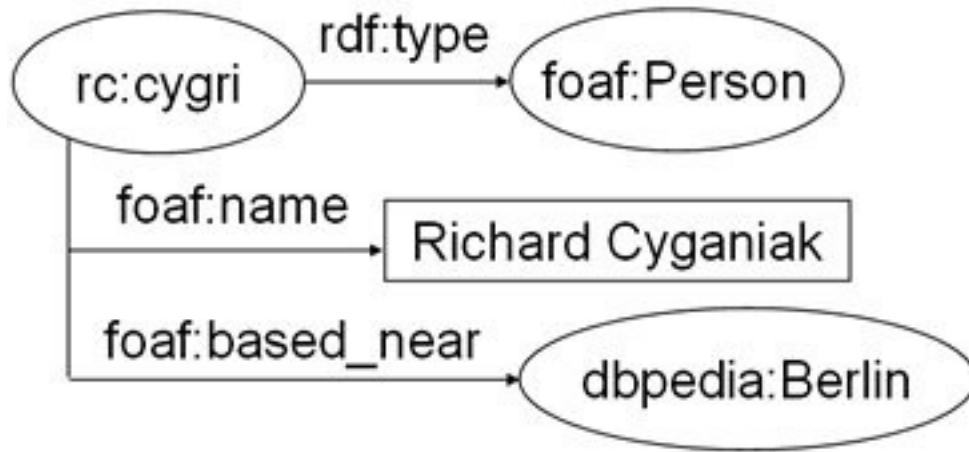
<http://sws.geonames.org/3333125/>

Data Merging with RDF

- Mix schemas/vocabularies within one document
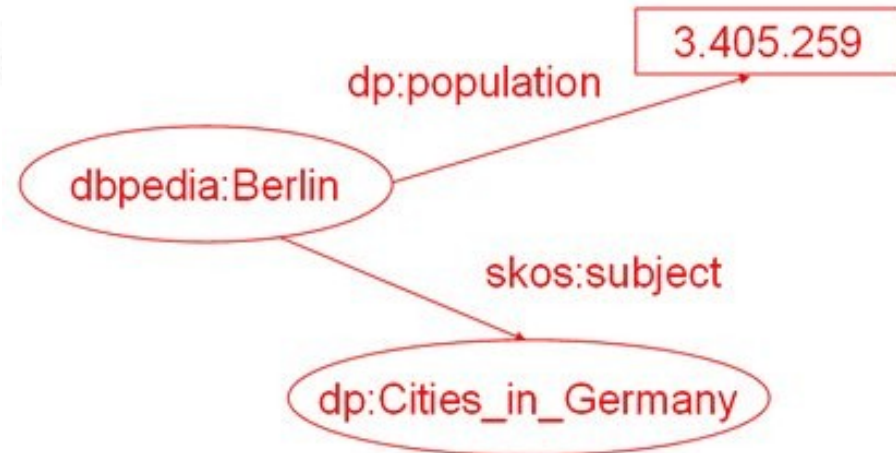
- Less painful data merging

Data Merging with RDF

Prefixes
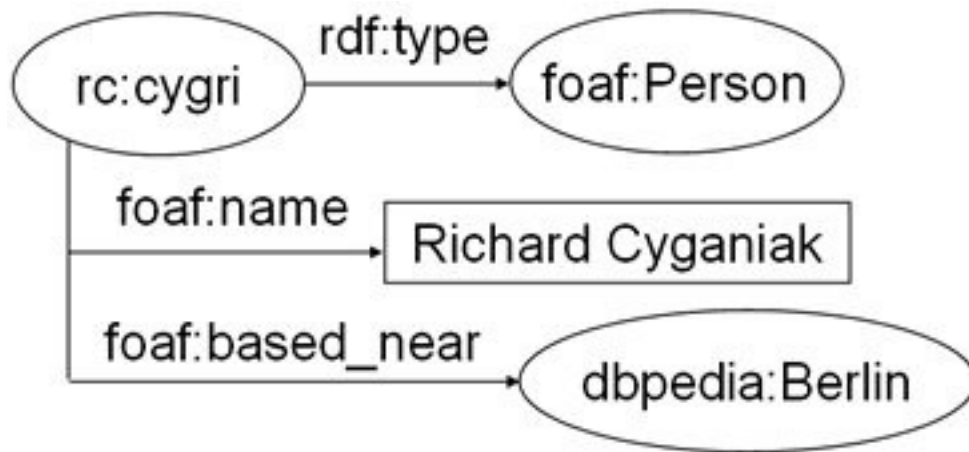rc: <http://richard.cyganiak.de/foaf.rdf#>
rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
foaf: <http://xmlns.com/foaf/0.1/>
dbpedia: <http://dbpedia.org/resource/>
dp: <http://dbpedia.org/property/>
skos: <http://www.w3.org/2004/02/skos/core#>

Data Merging with RDF

Prefixes
rc: <http://richard.cyganiak.de/foaf.rdf#>
rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
foaf: <http://xmlns.com/foaf/0.1/>
dbpedia: <http://dbpedia.org/resource/>
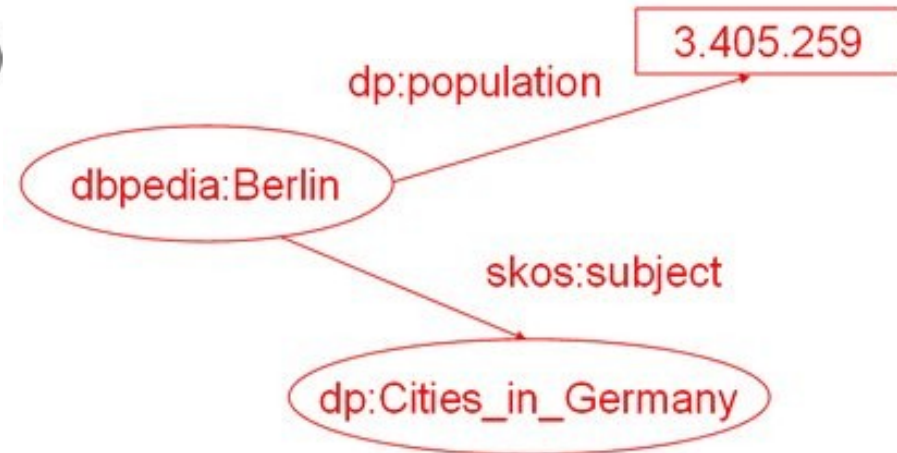dp: <http://dbpedia.org/property/>
skos: <http://www.w3.org/2004/02/skos/core#>
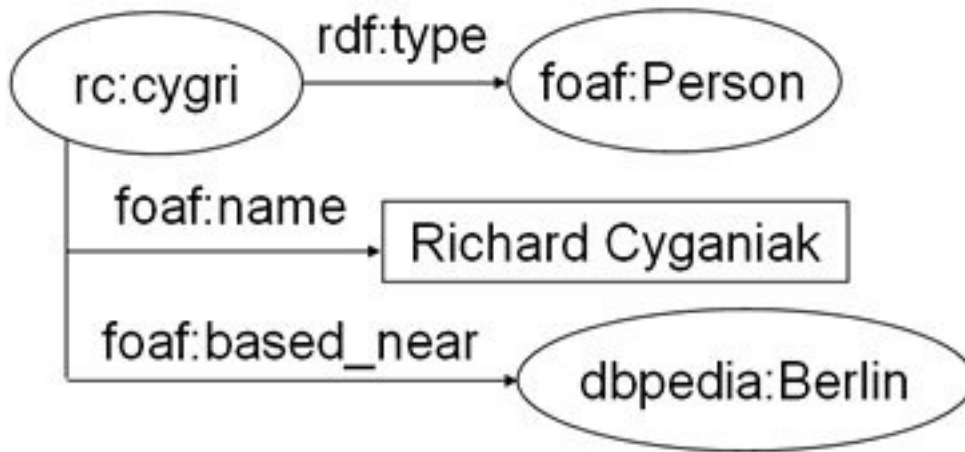
This is Linked Data

Linked Data Principles

1. Use URIs as names for things.

2. Use HTTP URIs so that people can look up those names.

3. When someone looks up a URI, provide useful RDF information.

4. Include RDF statements that link to other URIs so that they can discover related things.

Tim Berners-Lee 2007

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

Linked Data Benefits

- Reduces redundancy

- Enables (relatively) seamless data integration

- Allows network effects to add value to data

The Linking Open Data Project

The Linking Open Data Project

- Community project with W3C support

- Take existing open data sets

- Make them available on the Web in RDF

- Interlink them with other data sets

Participants

- Massachusetts Institute of Technology (US)
- University of Southampton (UK)
- Freie Universität Berlin (DE)
- DERI (IE)
- KMi, Open University (UK)
- University of London (UK)
- Universität Hannover (DE)
- University of Pennsylvania (US)
- Universität Leipzig (DE)
- Universität Karlsruhe (DE)
- Joanneum (AT)
- University of Toronto (CA)

- BBC (UK)
- Talis (UK)
- Garlik (UK)
- OpenLink (UK)
- Zitgist (US)
- Mondeca (FR)
- Cyc Foundation (US)

The LOD "Cloud" - May 2007

Geonames

talis®

GeoNames Home | Postal Codes | Download / Webservice | About   search [   ]   login

Map center : N 52° 28' 48'' W 1° 53' 40''   ExploreOurPla.net   google earth   tagzania   mapquest

| Map | Satellite | **Hybrid** | Terrain |

GeoNames Wikipedia

features
- ☑ city, village,...
- ☑ mountain,hill,rock,...
- ☑ stream, lake, ...
- ☑ country, state, region,...
- ☑ parks,area, ...
- ☑ road, railroad
- ☑ spot, building, farm
- ☑ forest,heath,...
- ☑ undersea
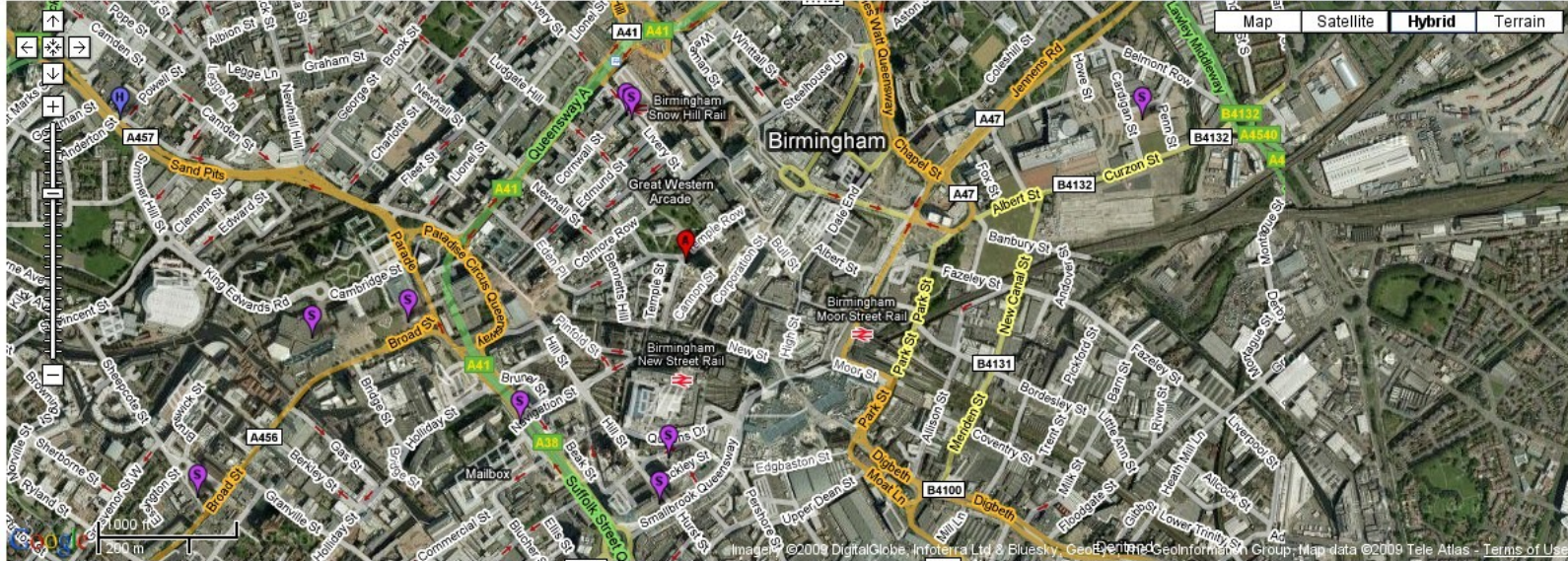
	Name	country	feature	km to center
1	City and Borough of Birmingham	United Kingdom	third-order administrative division	0.26 km
2	Comfort Inn Birmingham	United Kingdom	hotel	0.49 km
3	New Street Station	United Kingdom	railroad station	0.51 km
4	Snow Hill Station	United Kingdom	railroad station	0.51 km
5	Holiday Inn Birmingham City Centre	United Kingdom	hotel	0.58 km
6	Malmaison Birmingham	United Kingdom	hotel	0.7 km
7	Moor Street Station	United Kingdom	railroad station	0.84 km
8	Nitenite Birmingham Ltd	United Kingdom	hotel	0.87 km
9	The International Convention Centre	United Kingdom	community center	1.09 km
10	Brasshouse	United Kingdom	school	1.41 km
11	Worcester and Birmingham Canal	United Kingdom	canal	1.54 km

Birmingham

From Wikipedia, the free encyclopedia

Coordinates: 🌐 52°29′1″N 1°54′23″W

This article is about the city in England. For the U.S. city named after it, see Birmingham, Alabama. For other uses, see Birmingham (disambiguation).

Birmingham (🔊 pronunciation (help·info), /ˈbɜːmɪŋəm/ *BIRR-ming-əm*) is a city and metropolitan borough in the West Midlands county of England. Birmingham is the most populous of England's core cities, and is the second most populated British city after London, with a population of 1,010,200 (2005 estimate). Often considered to be the second city of the United Kingdom[2], the City of Birmingham forms part of the larger West Midlands conurbation, which has a population of 2,284,093 (2001 census)[3] and includes several neighbouring towns and cities, such as Solihull, Wolverhampton and the towns of the Black Country.

The city's reputation was forged as a powerhouse of the Industrial Revolution in England, a fact which led to Birmingham being known as "the workshop of the world" or the "city of a thousand trades".[4] Although Birmingham's industrial importance has declined, it has developed into a national commercial centre, being named as the third best place in the United Kingdom to locate a business, and the 21st best in Europe by Cushman & Wakefield in 2007.[5] It is also the fourth most visited city by foreign visitors in the UK.[6] In 1998, Birmingham hosted the G8 summit at the International Convention Centre, on the site of Bingley Hall, the world's first purpose-built exhibition hall, and remains a popular location for conventions today along with the National Exhibition Centre in nearby Solihull.[7] In 2007, Birmingham was ranked as the 55th most livable city in the world, according to the Mercer Index of worldwide standards of living. [8]

People from Birmingham are known as 'Brummies', a term derived from the city's nickname of *Brum*. This comes in turn from the city's dialect name, Brummagem,[9] which is derived from one of the city's earlier names, 'Bromwicham'. There is a distinctive Brummie dialect and accent, both of which differ from the adjacent Black Country.

Contents [hide]

Birmingham

— City and Metropolitan borough —

City of Birmingham

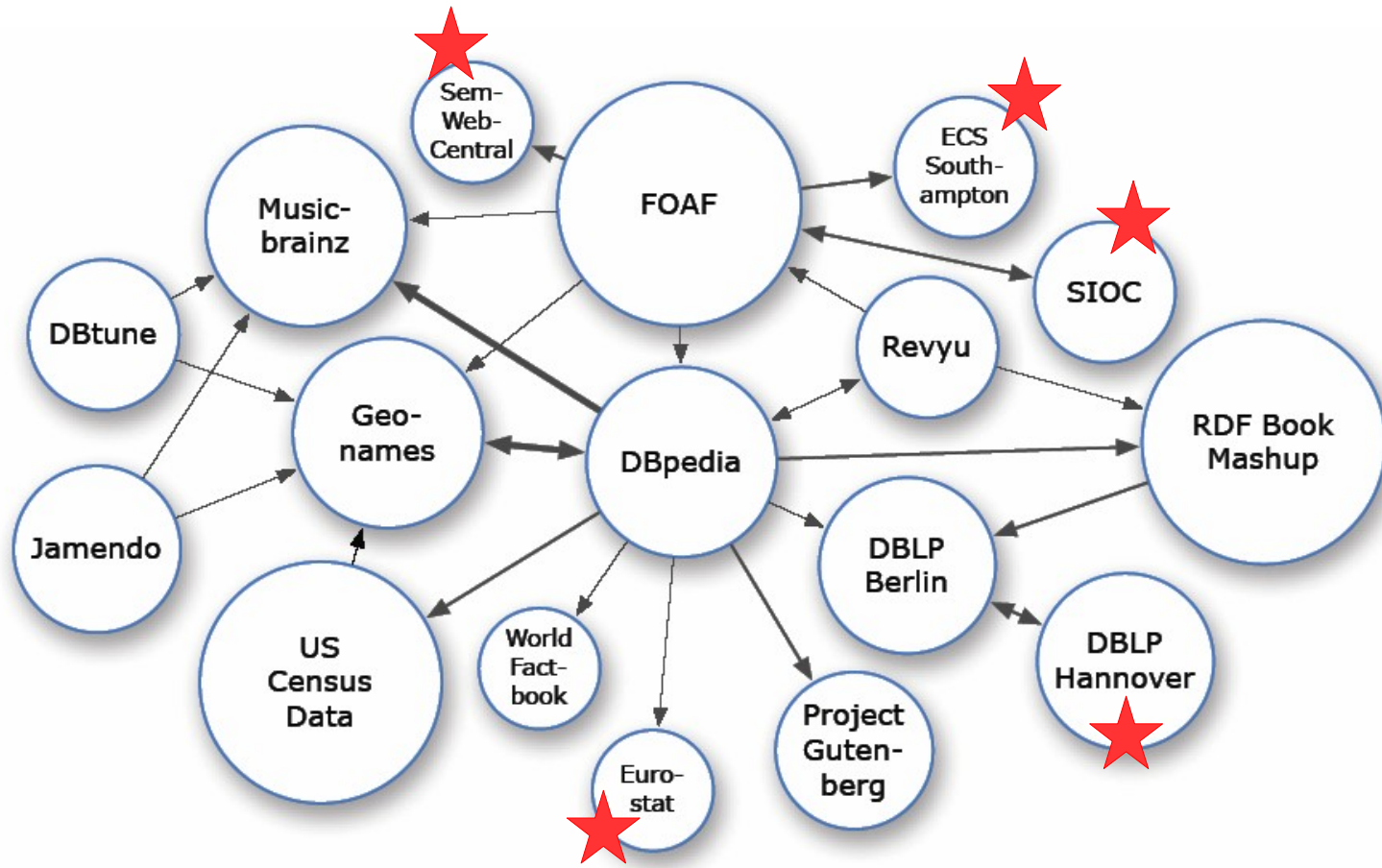Birmingham's skyline viewed from Bartley Green

Coat of Arms of the City Council

Nickname(s): "Brum", "Brummagem", "Second City", "Workshop of the World", "City of a Thousand Trades", "Canal City"
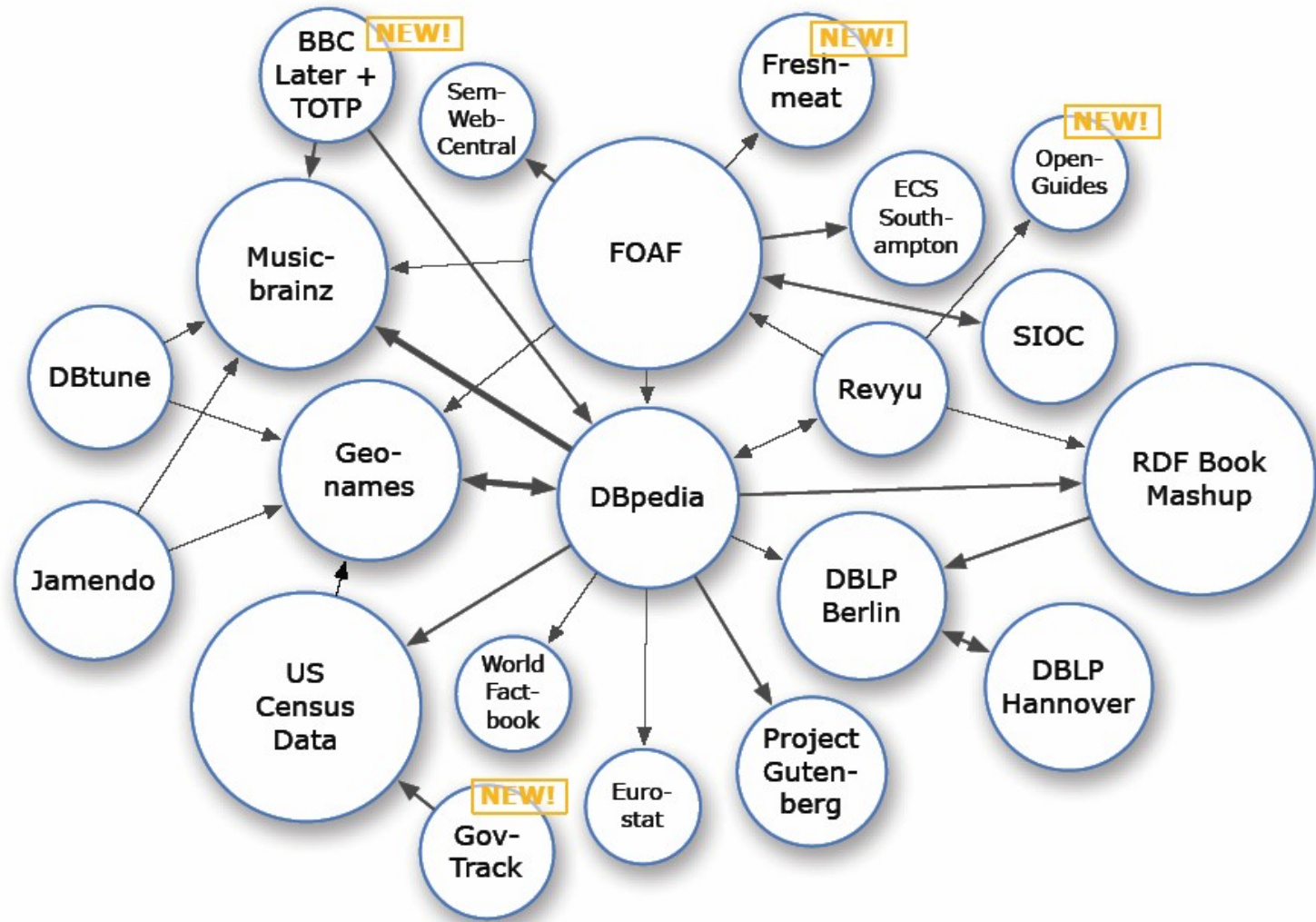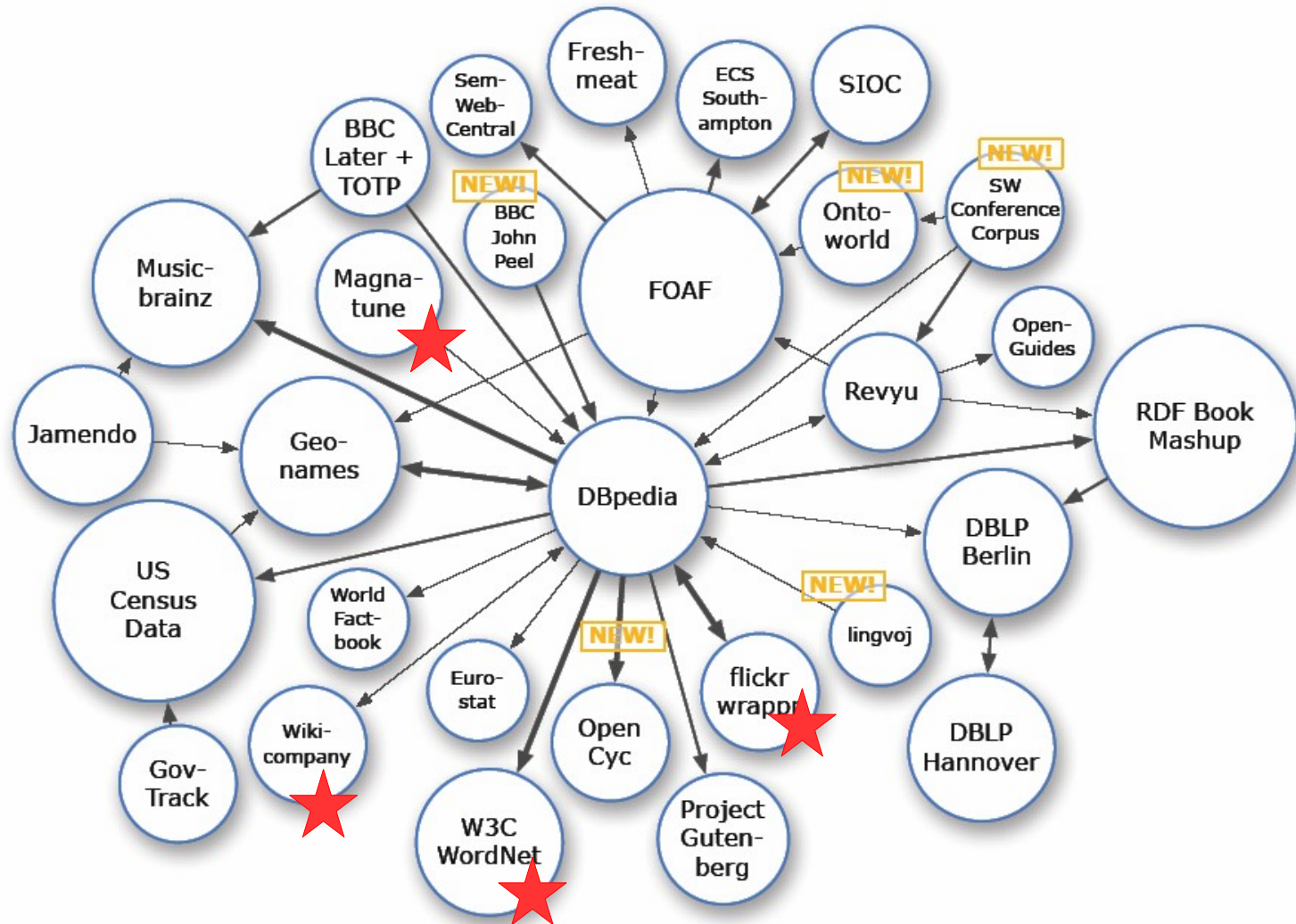
Motto: Forward

The LOD "Cloud" - July 2007

The LOD "Cloud" - August 2007
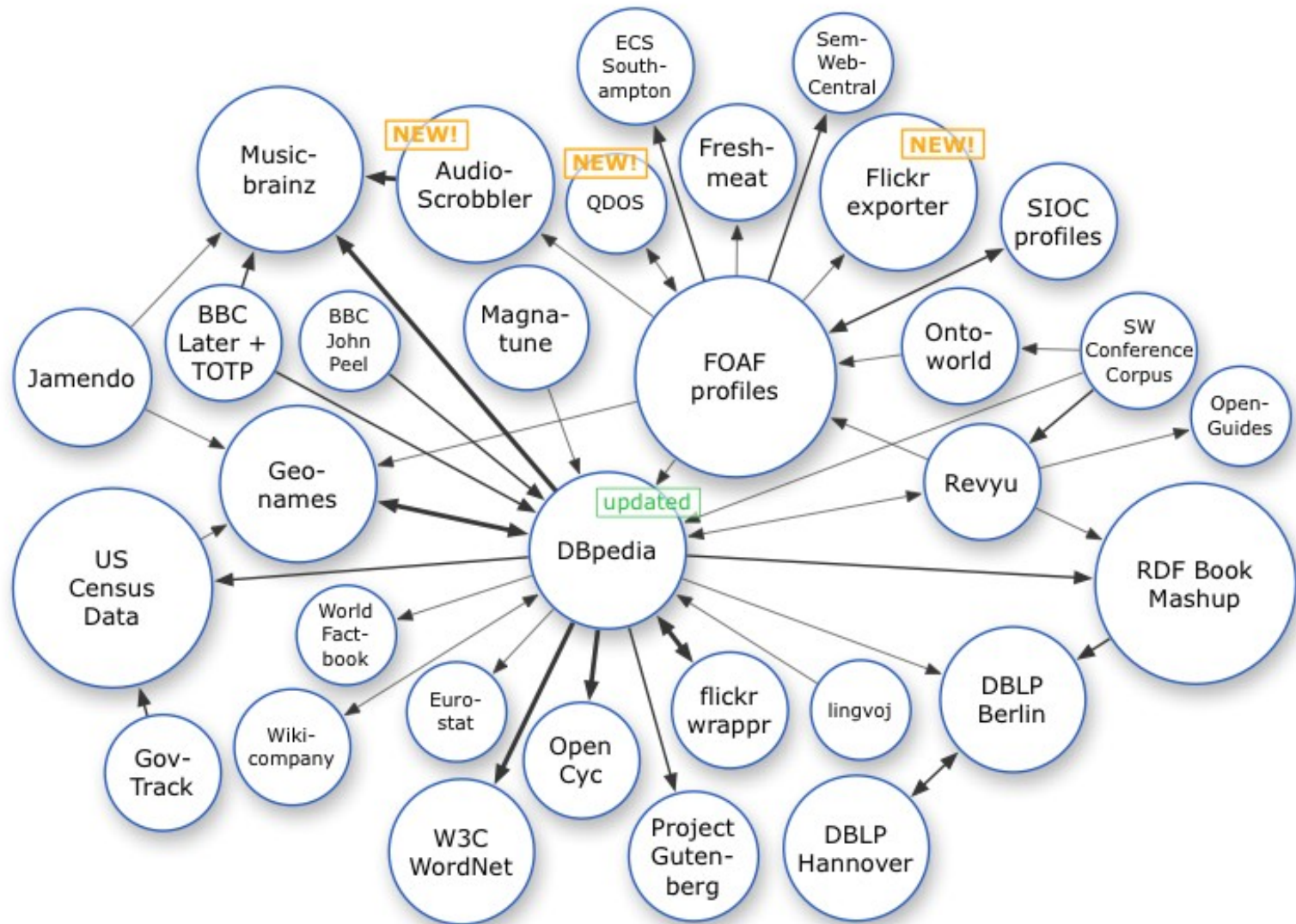
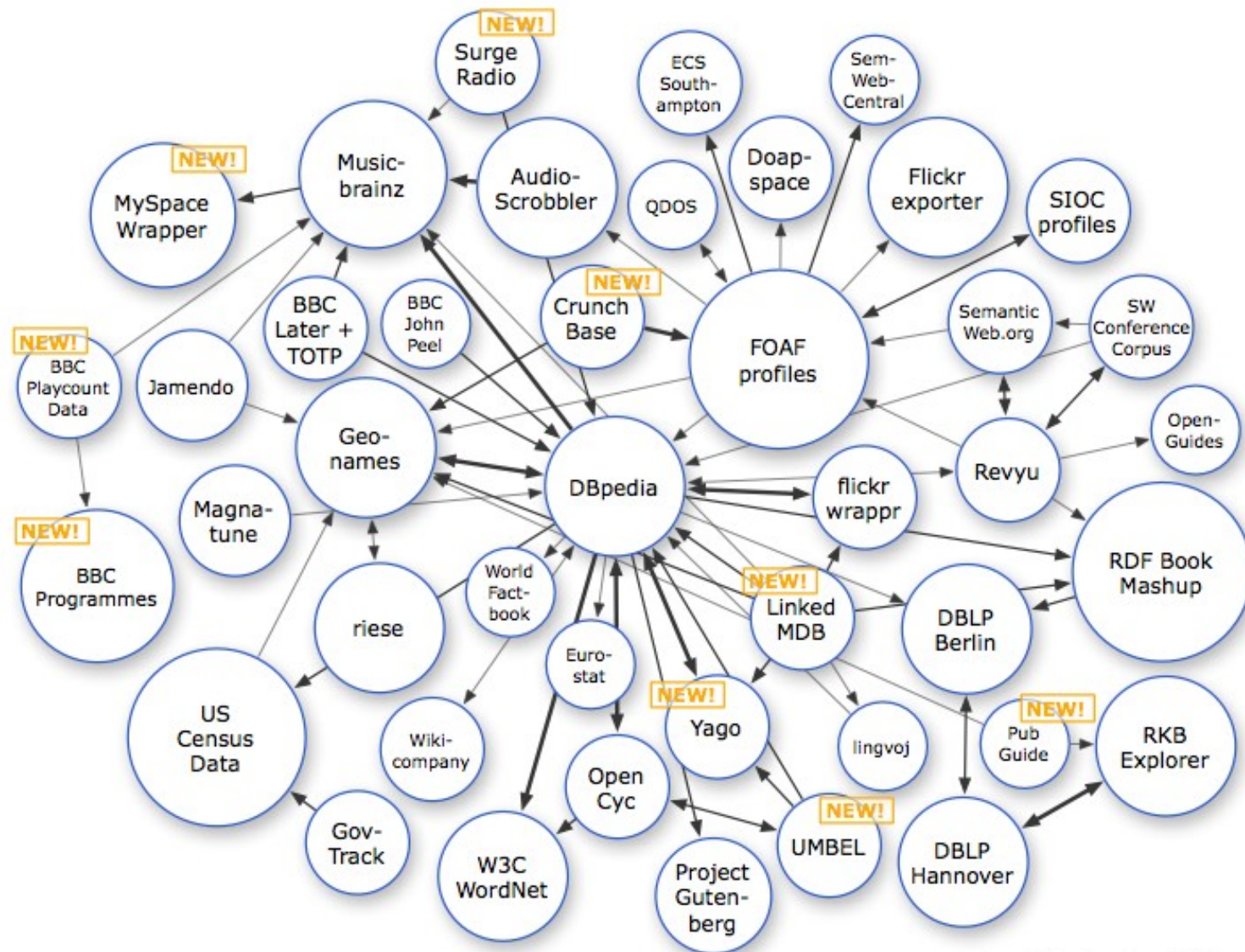# The LOD "Cloud" - November 2007

The LOD "Cloud" – Feb 2008

The LOD "Cloud" – Sept 2008

As of September 2008

Linked Data Applications

talis®

Linked Data Browsers

Linked Data Mashups

Search Engines

Thing

Thing

Thing

Thing

Thing

Thing

Thing

Thing

Thing

Thing

typed links

typed links

typed links

typed links

A

B

C

D

E
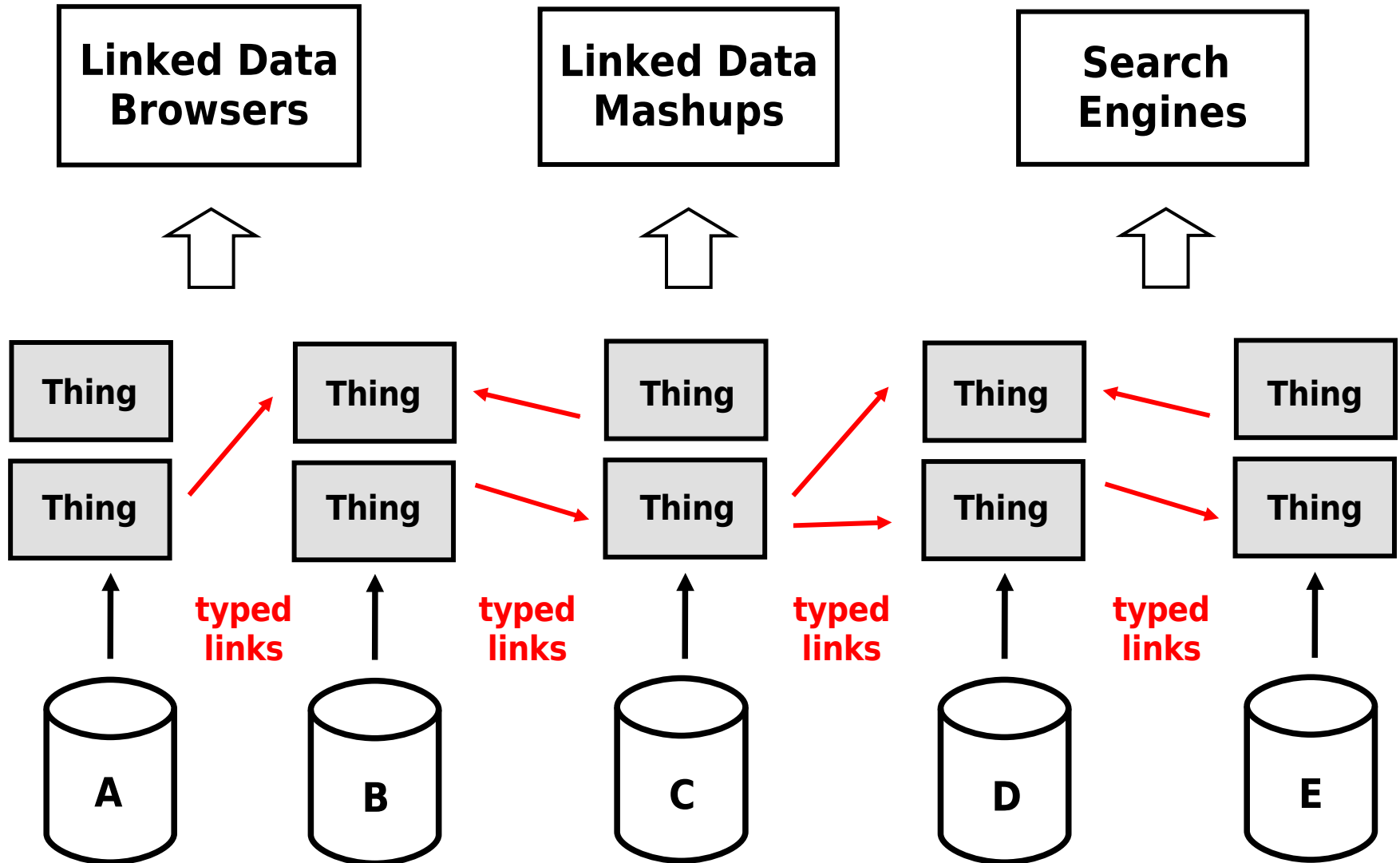
DBpedia Mobile

Revyu.com

Home | Browse Things | Search Things | Browse People

Login/Register | New Review

The Prestige

Links

Homepage: http://theprestige.movies.go.com/
See Also: http://imdb.com/title/tt0482571/

Tags

christian-bale christopher-nolan drama entertainment film hugh-jackman illusion magic michael-caine movie murder period scarlett-johansson science-fiction whodunnit

Reviews (1)

★★★★★ by martinp on 23 Jan 2007

This is a drama about intense rivalry between stage magicians in the late 19th Century. The evocation of the period, although first rate, is not the main attraction, however. The Prestige has an incredibly clever plot including the most ingenious murder I've ever come across. It also has a deeply moving and sad love story hidden in it, which gradually emerges over the course of the film.

The film requires a strong suspension of disbelief on some key points: there is a science-fiction premise which is introduced using the real historical character of Nikola Tesla (I'd rather they had used a fictional scientist). There are a couple more implausibilities required to hold it together (something odd that goes on that none of the characters pick up on and a dead-end that by a huge coincidence turns out not to be a dead-end: I can't be more specific without spoiling the plot).

However, rather than feeling cheated by these aspects of the film, I'm hugely impressed. The writers have taken an implausible (okay, impossible) premise but created an intricate, involving and visual story that would be impossible without that premise. Scenes join up with each other in many subtle ways, echoing the same writers' earlier film Memento. Even when you've seen the twist coming, the final scene which lays it all out are has a lot of impact and I suspect the final shot will haunt my dreams.

I expected the film to be about nice costumes or impressive magical trickery, but it is actually about deep emotions felt by the main characters as they deal with the situations life has dealt them, and it rather than serving up those emotions on a plate, it requires you to think and piece together what you've seen. That's got to be a good thing, in fact the best of what film a be.

What do you think of **The Prestige**? **Write Your Own Review...**

Revyu.com: Contact | Credits | Privacy Policy | Disclaimer

The Prestige

directed by **Christopher Nolan**

RDF Metadata About
The Prestige

RDF META

Write a Review of
The Prestige

Add to del.icio.us

Falcons

Falcons

Object   Concept   Document

birmingham   Search Objects

Separate keywords with a space, and put a phrase in double quotes.

Specify a type:

Area	Attraction	Body	Building	City
Concept	Landmark	**Location**	Medium	Museum
Organization	Person	Railway Station	Sport	Subject

Objects **1 - 10** of **9,177** for your search **birmingham** (**2.03** seconds)

Birmingham   is a   *Borough*, *City*
- abstract: :**Birmingham** i USA hittas på **Birmingham**, Alabama **Birmingham** ...   - From dbpedia.org »
- has subject: Category:1st_millennium_establishments   - From dbpedia.org »
- hasPhotoCollection: **Birmingham**   - From dbpedia.org »
http://dbpedia.org/resource/Birmingham - Described in 553 documents

People from **Birmingham**, Alabama   is a   *Concept*
☆ has broader: Category **Birmingham** 2 C Alabama   - From dbpedia.org »
☆ preferred label: People from **Birmingham**, Alabama   - From dbpedia.org »
☆ label: People from **Birmingham**, Alabama   - From dbpedia.org »
http://dbpedia.org/resource/Category:People_from_Birmingham%2C_Alabama - Described in 255 documents

Birmingham_City_F.C.   is a   *Club*, *Company*
- abstract: ...-cn:伯明翰足球乐部}-{zh-hk:伯明翰城足球會;zh-cn:伯明翰足球俱乐部}- (**Birmingham** City Football Club)是一家英格蘭-{zh-hk:球會;zh-cn:俱樂部}-...   - From dbpedia.org »
- has subject: Category:English_football_clubs   - From dbpedia.org »
- hasPhotoCollection: **Birmingham** City F C   - From dbpedia.org »
http://dbpedia.org/resource/Birmingham_City_F.C. - Described in 205 documents

Birmingham   is a   *Ontology*, *Subject*, *City*, *Geographical region*
☆ imports: http://semantic-mediawiki.org/swivt/1.0   - From semanticweb.org »
☆ date of creation: 2008-05-29T13:42:38+02:00   - From semanticweb.org »
☆ page: **Birmingham**   - From semanticweb.org »
http://semanticweb.org/id/Birmingham - Described in 10 documents

Birmingham%2C_Alabama   is a   *City*
- abstract: :**Birmingham** to także nazwa miasta w Anglii. **Birmingham** - mi...   - From dbpedia.org »
- has subject: Category:United_States_communities_with_African_American_majority_populations   - From dbpedia.org »
- hasPhotoCollection: **Birmingham** 2 C Alabama   - From dbpedia.org »
http://dbpedia.org/resource/Birmingham%2C_Alabama - Described in 238 documents

Birmingham_Barons   is a   *Sport*, *Team*
- abstract: The **Birmingham** Barons are a minor league baseball team based i...   - From dbpedia.org »
- has subject: Category:Minor_league_baseball_teams   - From dbpedia.org »
- hasPhotoCollection: **Birmingham** Barons   - From dbpedia.org »
http://dbpedia.org/resource/Birmingham_Barons - Described in 13 documents

Ongoing Challenges

Linking Algorithms and Data Fusion

- How can we automatically and reliably link related items in distributed data sets?
  - string matching
  - common keys
  - shared properties
  - ?

- How can we automatically select data sets for fusion?

Attribution and Licensing in the Web of Data

- When data can be easily integrated from multiple sources, how do we communicate the source and licensing conditions of each each fragment?

Understanding New Business Models

- When data can be easily integrated from multiple sources, how do each of the individual data providers generate revenue from their data?

Enabling New User Interactions

- When data can be easily integrated from multiple sources, what new forms of user interaction may be possible?

Linked Data Outlook in 2009

- Linked Data on the Web workshop at WWW2009
- Special Issue of IJSWIS on Linked Data
- More data sets
- More applications
- More activities worldwide
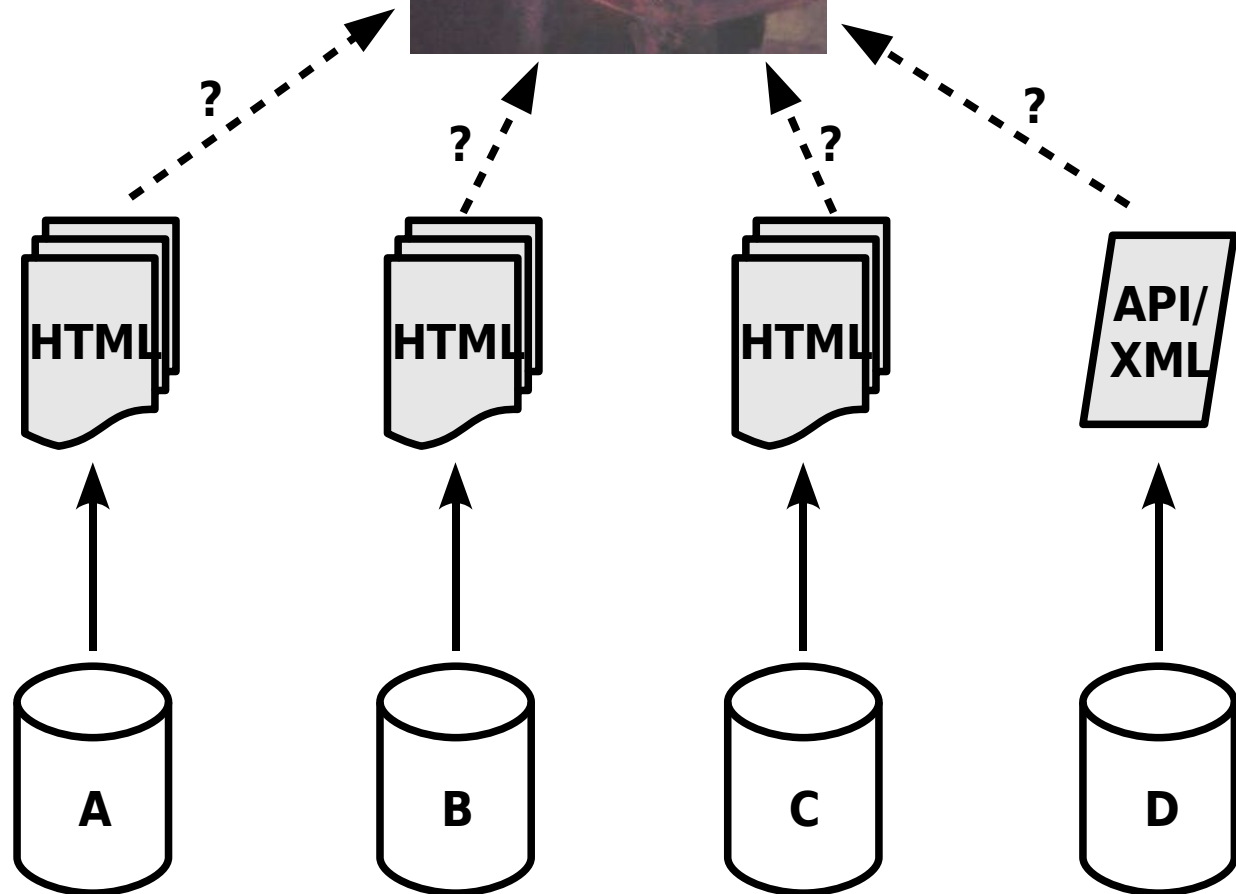- Linked Data tutorials here in Austin, tomorrow and Saturday!

Questions?

- Contact Details
  - http://tomheath.com
  - http://www.talis.com
  - tom.heath@talis.com

- Further Information
  - http://linkeddata.org

shared innovation

How do you know these documents are about Rembrandt?

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