

**17th International World Wide Web Conference
W3C Track @ WWW2008, Beijing, China
23-24 April 2008**

Linked Data: Principles and State of the Art

**Christian Bizer, Freie Universität Berlin
Tom Heath, Talis
Tim Berners-Lee, W3C/MIT**



1. The Web of Documents and the Web of Data

- **From global filesystem to global database**

2. The W3C SWEO Linking Open Data Project

- **Bootstrapping the Web of Linked Data**

3. What is next?

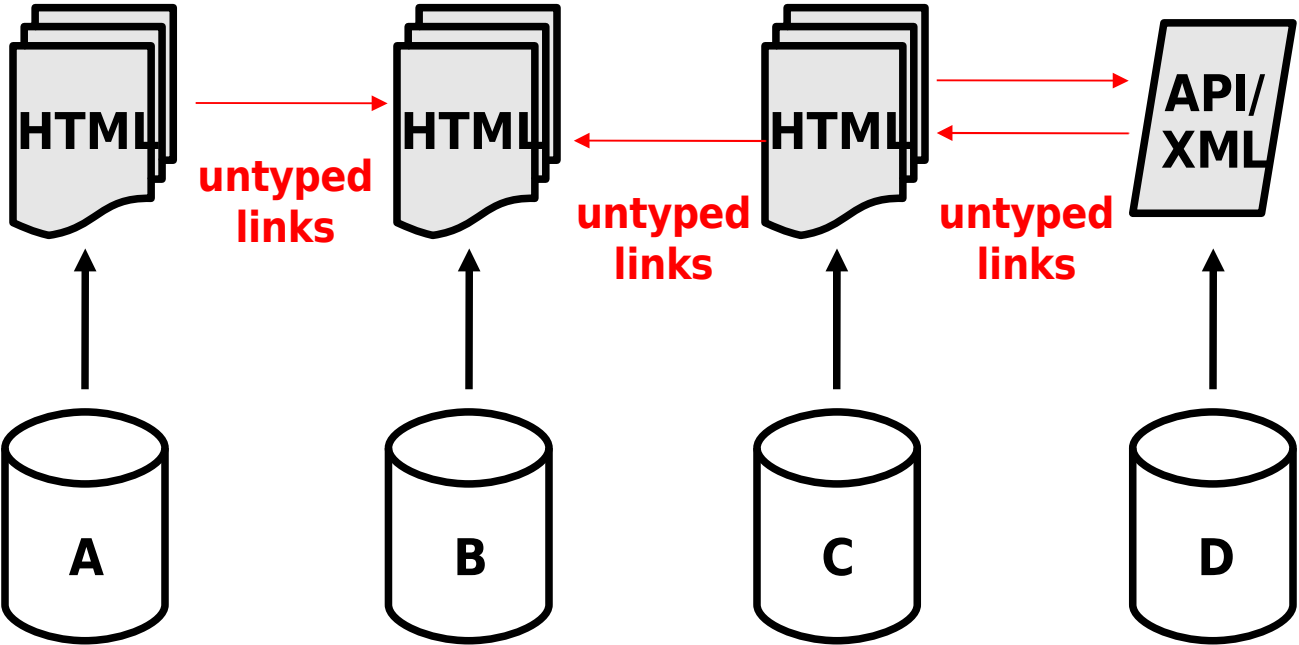
- **Open Issues and directions for future work**

The Web of Documents

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



The Web of Documents



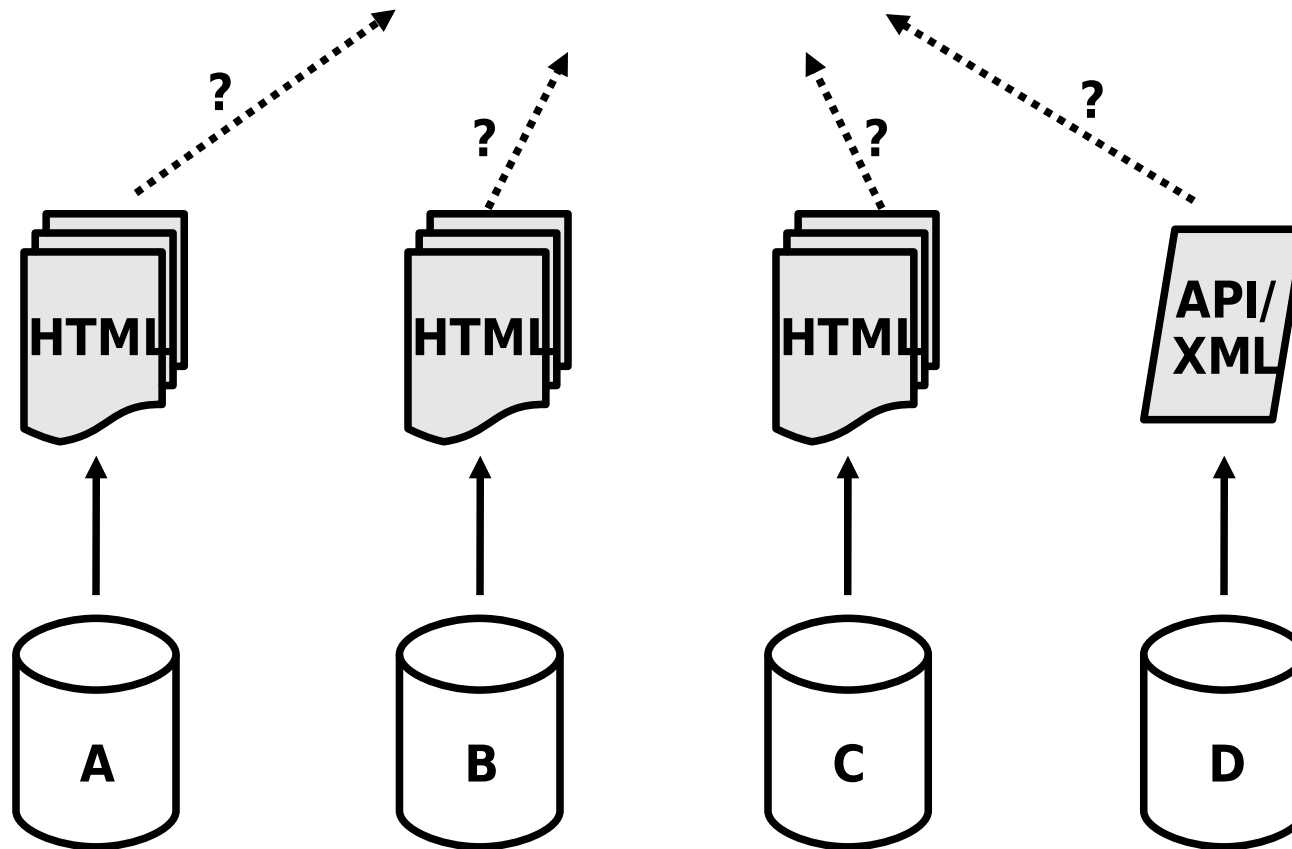
Simplicity



Data Silos on the Web



How do you know these documents are about Beijing?



Disconnected Data



The Web of Documents: Challenges

■ Data Integration

- Show me all the publications from Semantic Web-related conferences in 2007

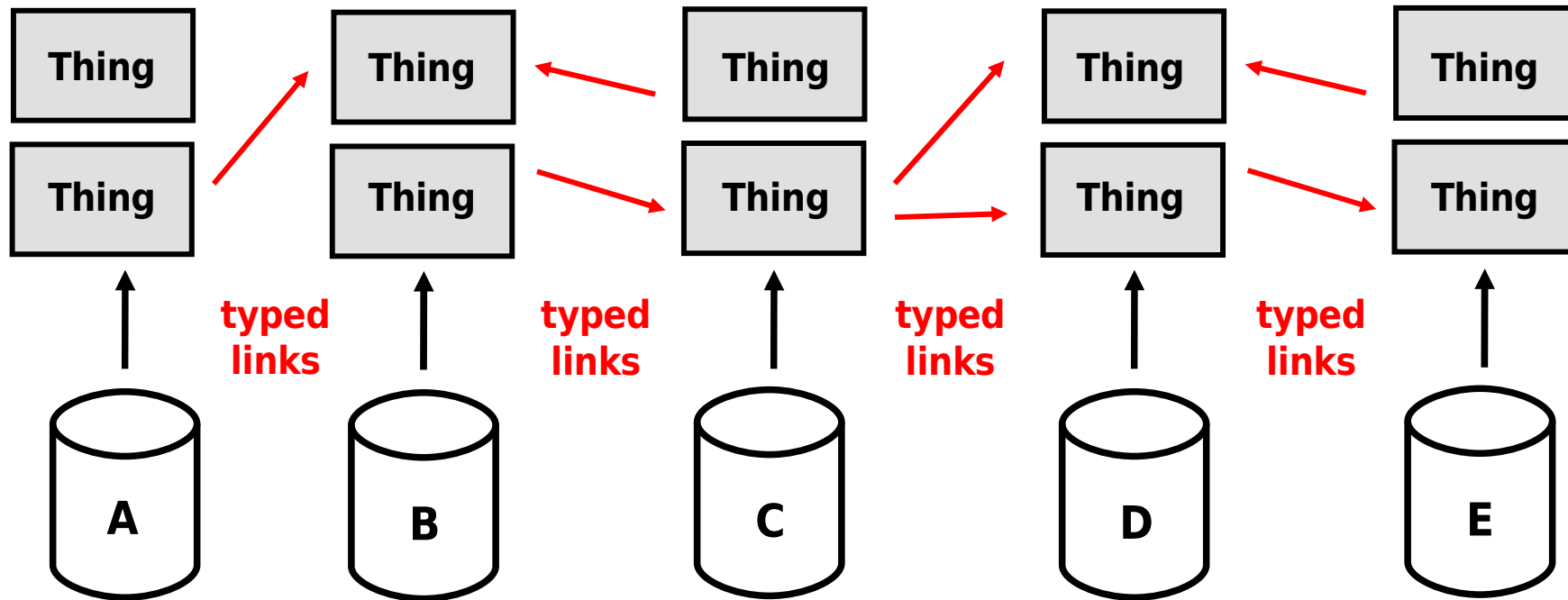
■ Querying Across Data Sources

- Which WWW2008 papers have been written by people from companies of less than 100 people?

Linked Data

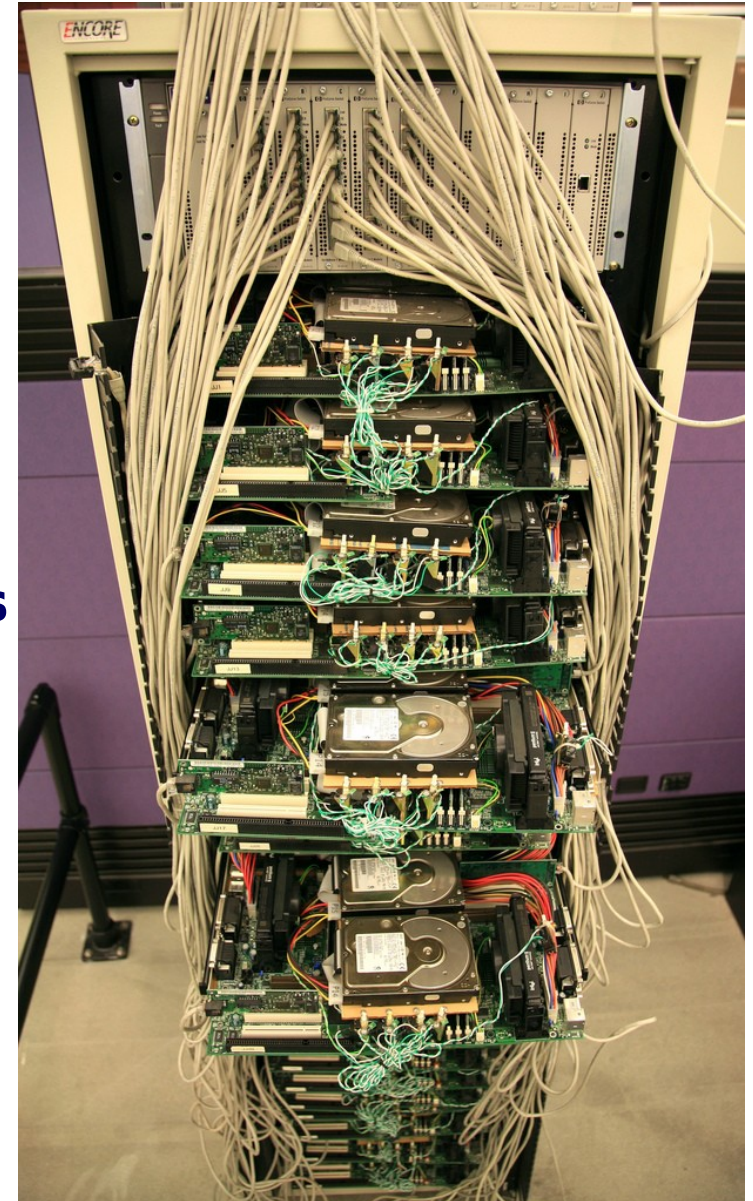
- **Many common things are represented in multiple data sets**
- **Linking identifiers connects these data sets**
- **Linked data opens the doors of the silos**

Linked Data



The Web of Data

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



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>

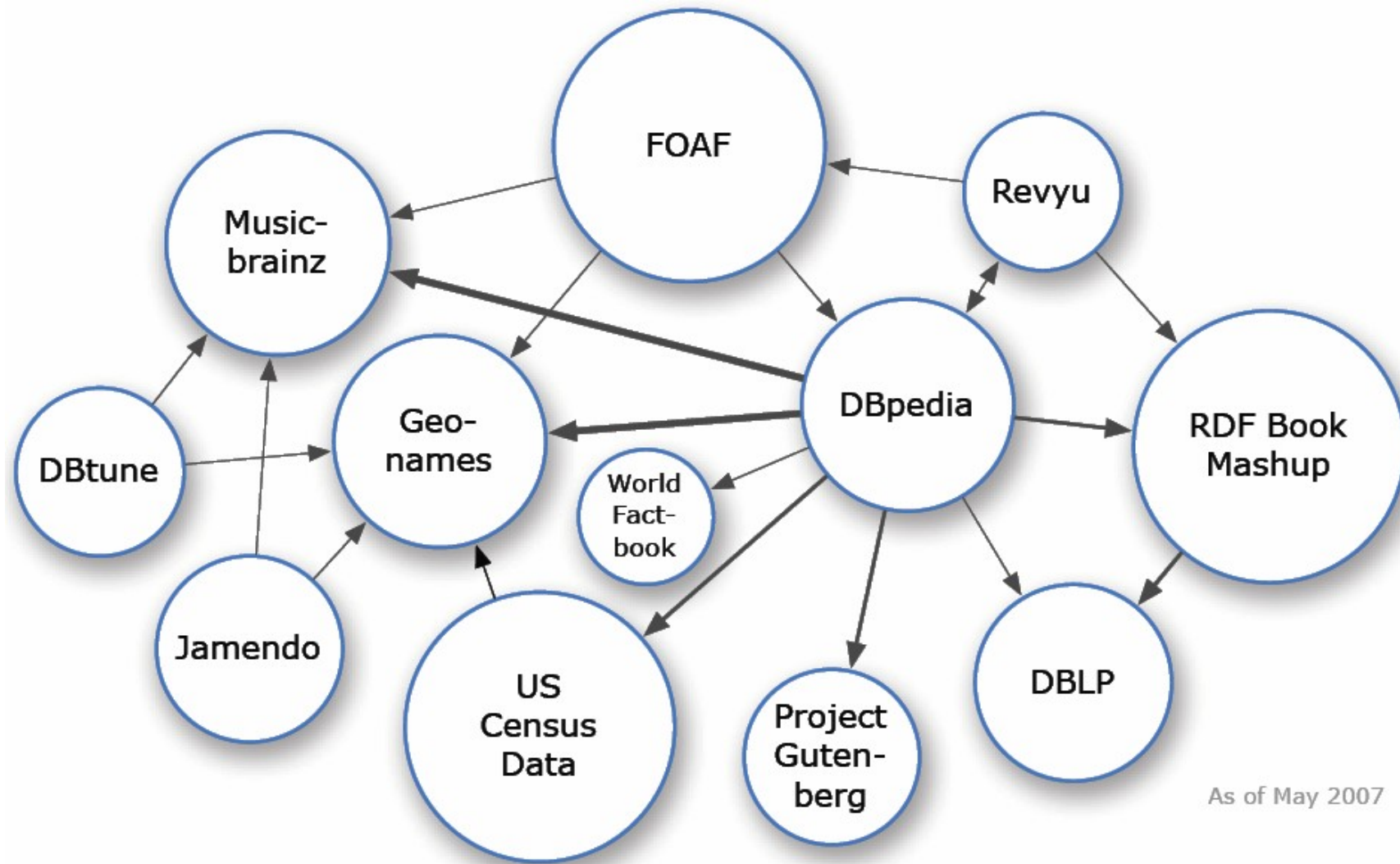
2. W3C SWEO Linking Open Data Project



Community effort to

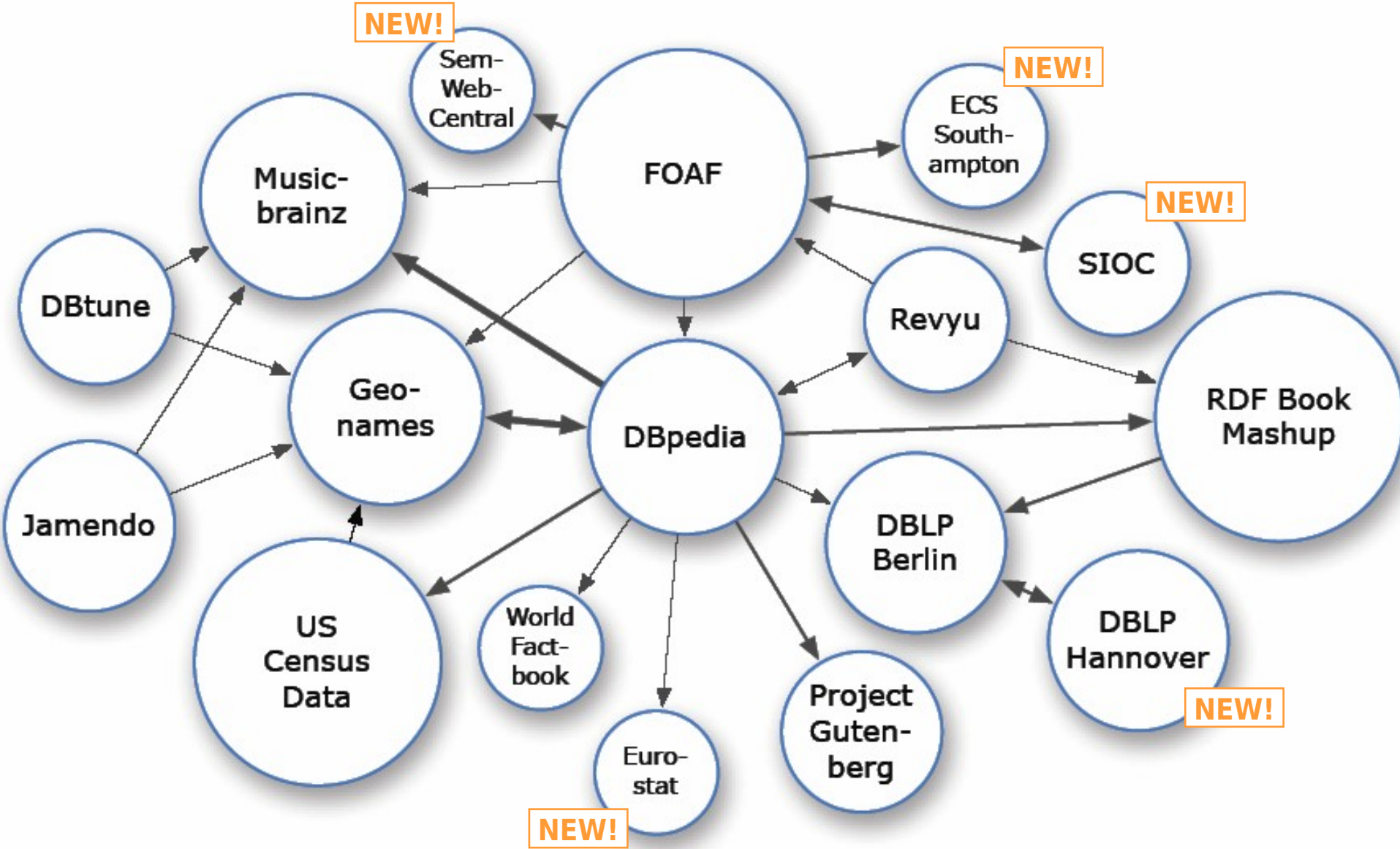
- **publish existing open license datasets as Linked Data on the Web**
- **interlink things between different data sources**
- **develop clients that consume Linked Data from the Web**

LOD Datasets on the Web: May 2007

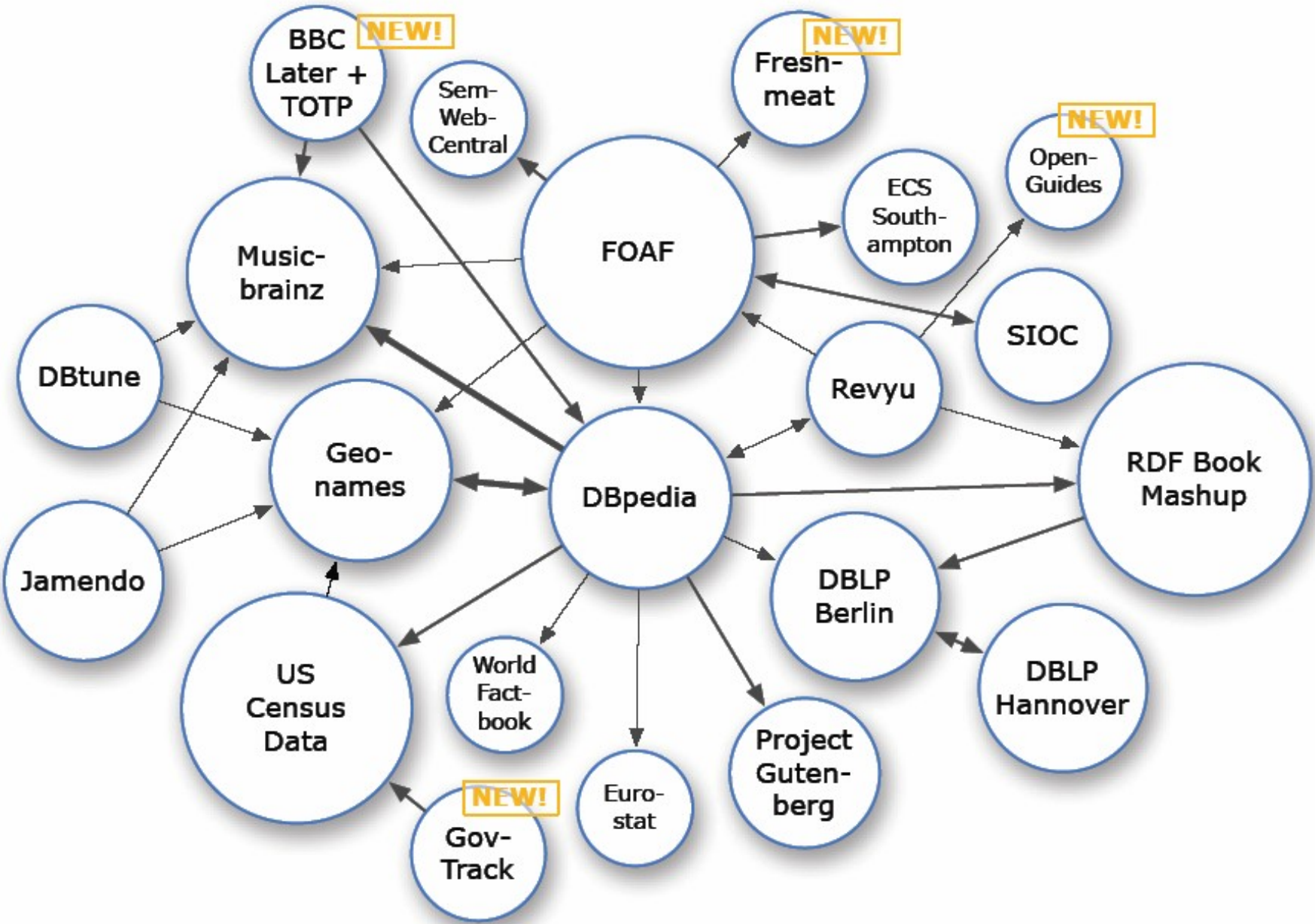


- Over 500 million RDF triples
- Around 120,000 RDF links between data sources

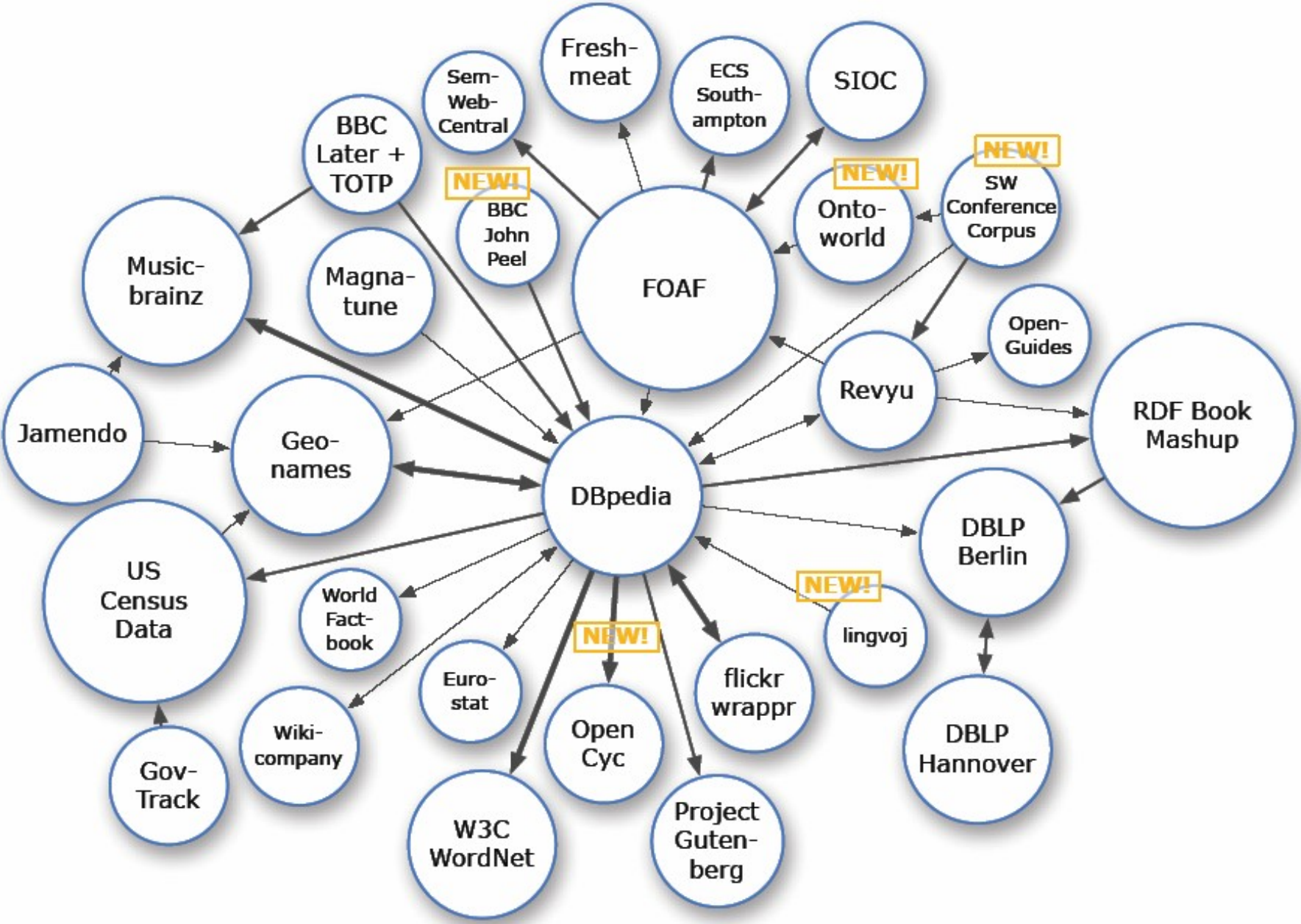
LOD Datasets on the Web: July 2007



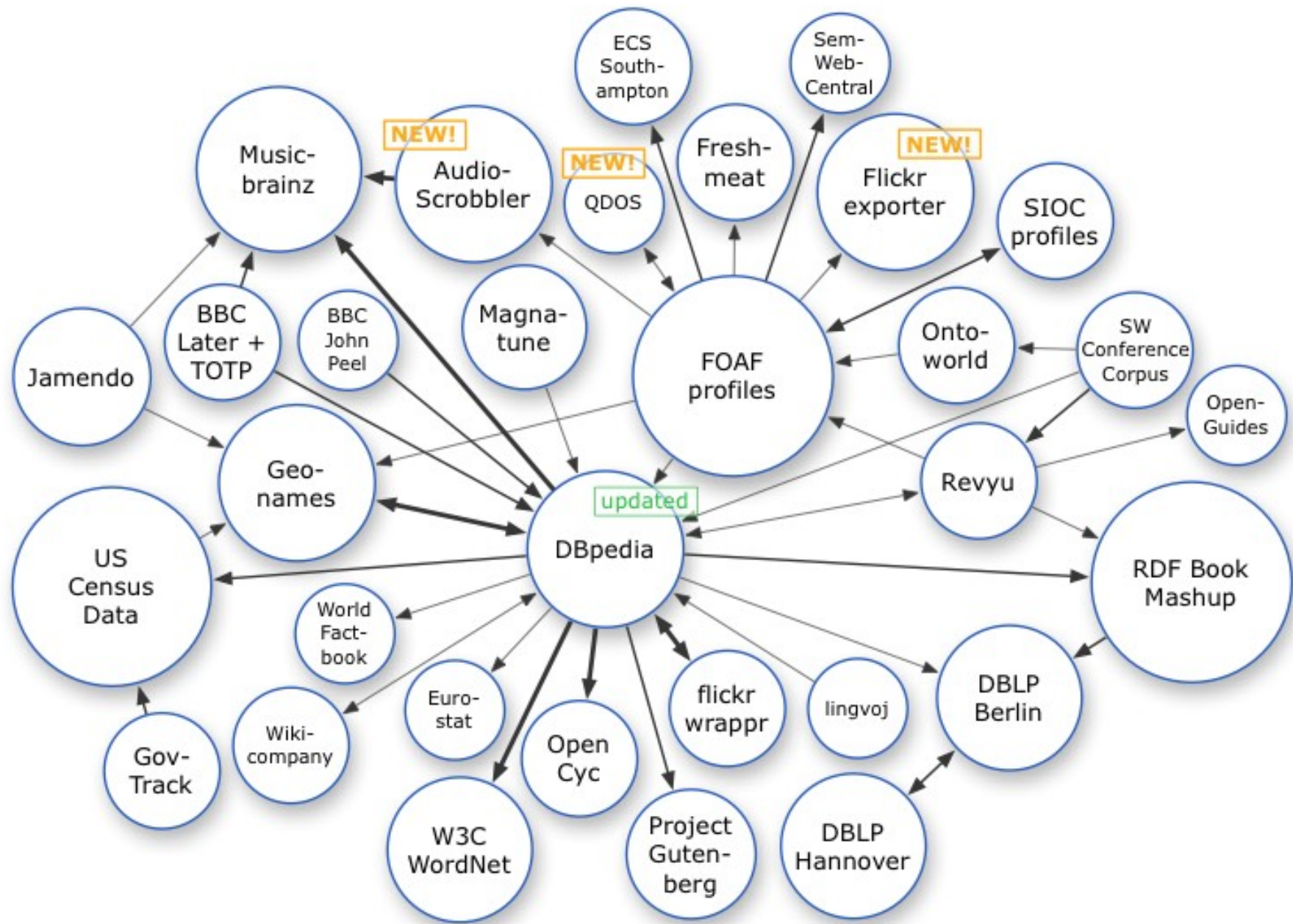
LOD Datasets on the Web: August 2007



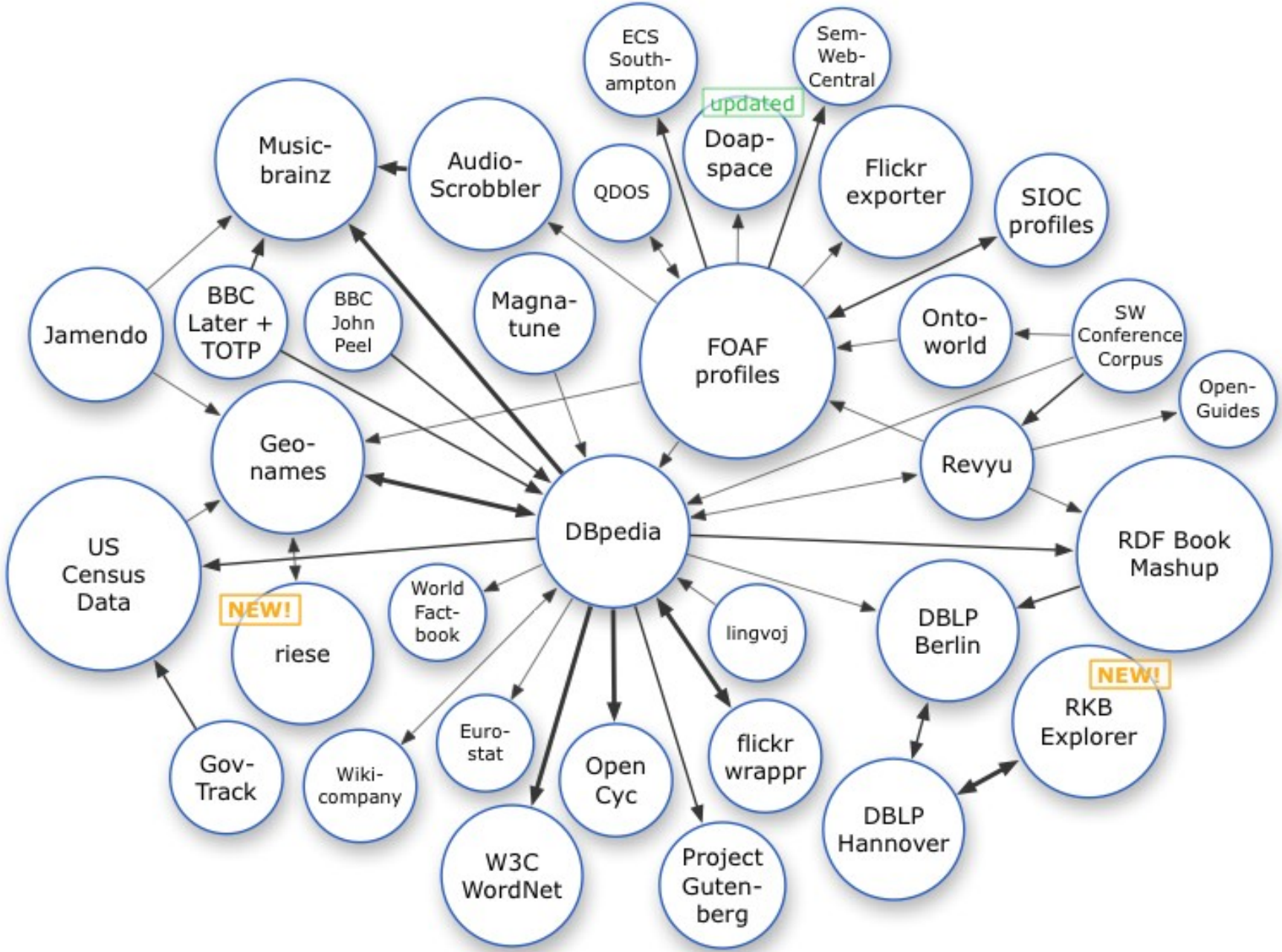
LOD Datasets on the Web: November 2007



LOD Datasets on the Web: February 2008



LOD Datasets on the Web: April 2008



Triple Count

- **More than 2 billion RDF triples**
- **Interlinked by around 3 million RDF links**

(rough estimates)

Organizations participating in the LOD community

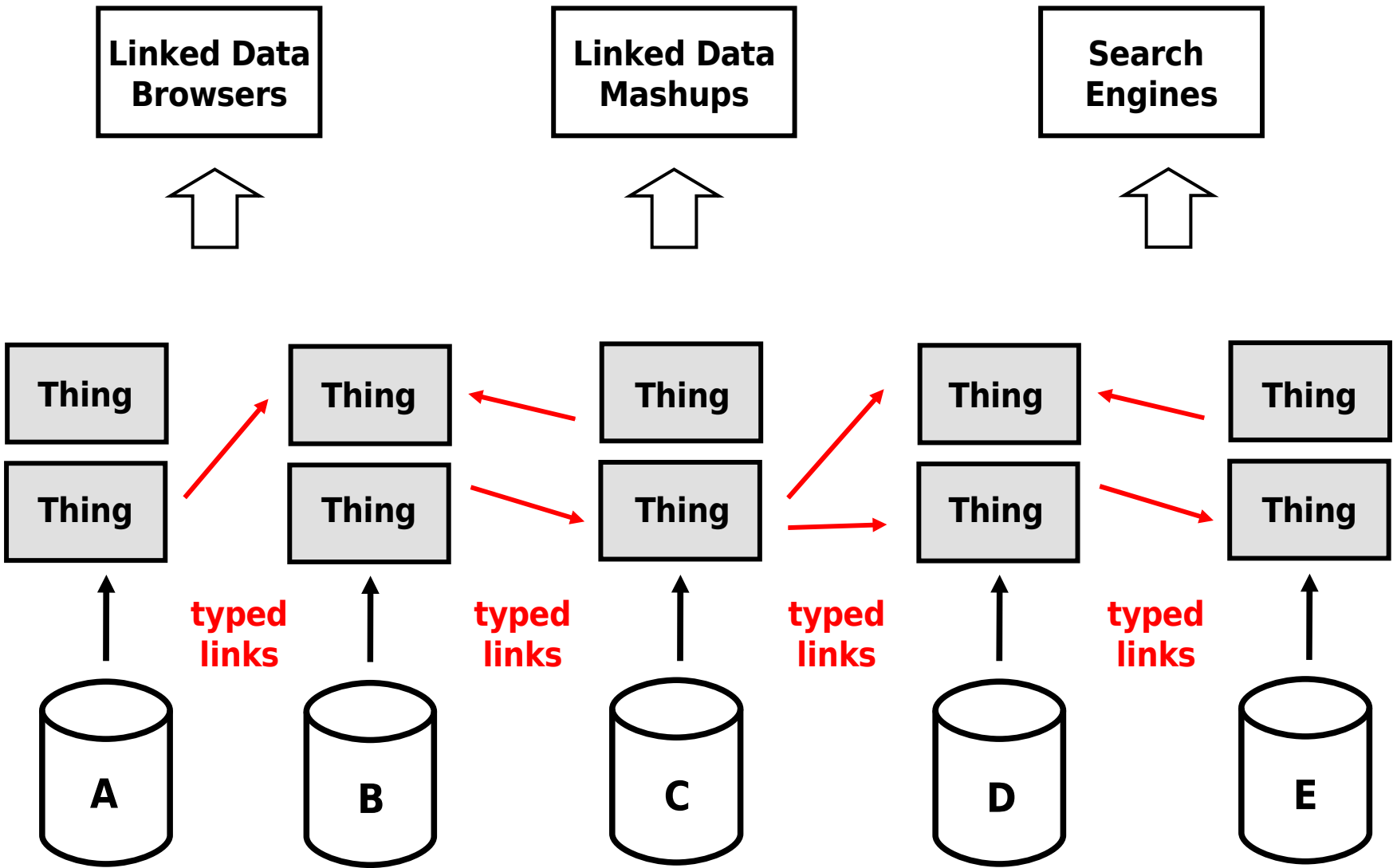
■ Universities and Research Institutes

- **Massachusetts Institute of Technology (USA)**
- **University of Southampton (UK)**
- **DERI (IRE)**
- **KMi, Open University (UK)**
- **University of London (UK)**
- **Universität Hannover (DE)**
- **University of Pennsylvania (USA)**
- **Universität Leipzig (DE)**
- **Universität Karlsruhe (DE)**
- **Joanneum (AT)**
- **Freie Universität Berlin (DE)**
- **Cyc Foundation (USA)**
- **SouthEast University (CN)**

■ Companies

- **BBC (UK)**
- **OpenLink (UK)**
- **Talis (UK)**
- **Zitgist (USA)**
- **Garlik (UK)**
- **Mondeca (FR)**
- **Renault (FR)**
- **Boab Interactive (AUS)**

So what can we do with this?



Linked Data Browsers

- Tabulator Browser (MIT, USA)
- Disco Hyperdata Browser (FU Berlin, DE)
- OpenLink RDF Browser (OpenLink, UK)
- Zitgist RDF Browser (Zitgist, USA)
- Humboldt (HP Labs, UK)
- Fenfire (DERI, Ireland)
- Marbles (FU Berlin, DE)

Beijing Capital International Airport

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

- [Building](#) 
- [Feature](#) 

[comment](#)

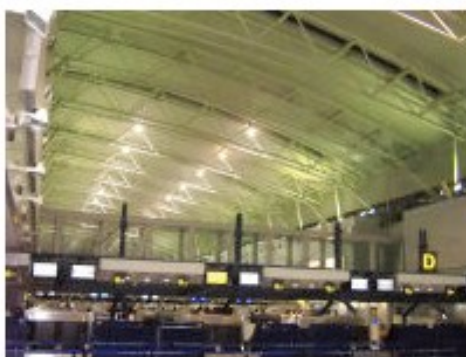
- Der Flughafen Peking ist einer der wichtigsten Flughäfen Chinas und liegt im Nordosten Pekings. Er liegt 20 Kilometer nordöstlich vom Stadtzentrum, im Gebiet Shunyi. Genauer ist aber der Flughafen innerhalb des Verwaltungsgebiets des Stadtbezirks Chaoyang. 
- L'aéroport international de Pékin (code AITA : PEK ; code OACI : ZBAA) est le principal aéroport de Pékin, capitale de la République populaire de Chine. Il est situé à une vingtaine de kilomètres au nord-est du centre de la ville, dans le district de

⋮

is sameAs of

- [Beijing Capital International Airport](#) 













[depiction](#)















⋮

This is release 0.8 of the Tabulator Project. The [live development trunk](#) is also available.

▼ Tim Berners-Lee

- request
 - ▶ Request for <<http://www4.wiwiss.fu-berlin.de/bookmashup/persons/Tim+Berners-Lee>>
 - ▶ Request for <<http://www4.wiwiss.fu-berlin.de/dblp/resource/person/100007>>
- requested by
 - ▶  Tim Berners-Lee's FOAF file 
- type
 - ▶  Male
 - ▶  Person 
- label
 - Tim Berners-Lee
- seeAlso
 - ▶  RDF document describing: Tim Berners-Lee 
 - ▶  RDF Description of Tim Berners-Lee 
- assistant
 - ▶  Amy van der Hiel
- homePage
 - ▶  <http://www.w3.org/People/Berners-Lee/> 
- work
 - ▶ ...
- organization
 - ▼ **World Wide Web Consortium**

- title
 - W3C (large: includes all specs and groups)
- type
 - ▶  SocialEntity
 - ▶  Organization 
- label
 - W3C
- seeAlso
 - ▶  W3C Groups and Organizational Structure 
 - ▶  W3C Standards and Technical Reports 
- publicHomePage
 - ▶  <http://www.w3.org/> 
- homepage
 - ▶  <http://www.w3.org/> 
- logo
 - ▶ 
- name
 - World Wide Web Consortium
- is organization of
 - ▶  Tim Berners-Lee

- publicHomePage
 - ▶  <http://www.w3.org/People/Berners-Lee/> 

- ▶  Tim Berners-Lee

- **Domain-specific applications using Linked Data from the Web**

- Website for rating everything
- Uses DBpedia data to augment ratings

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](#)

DBpedia Mobile



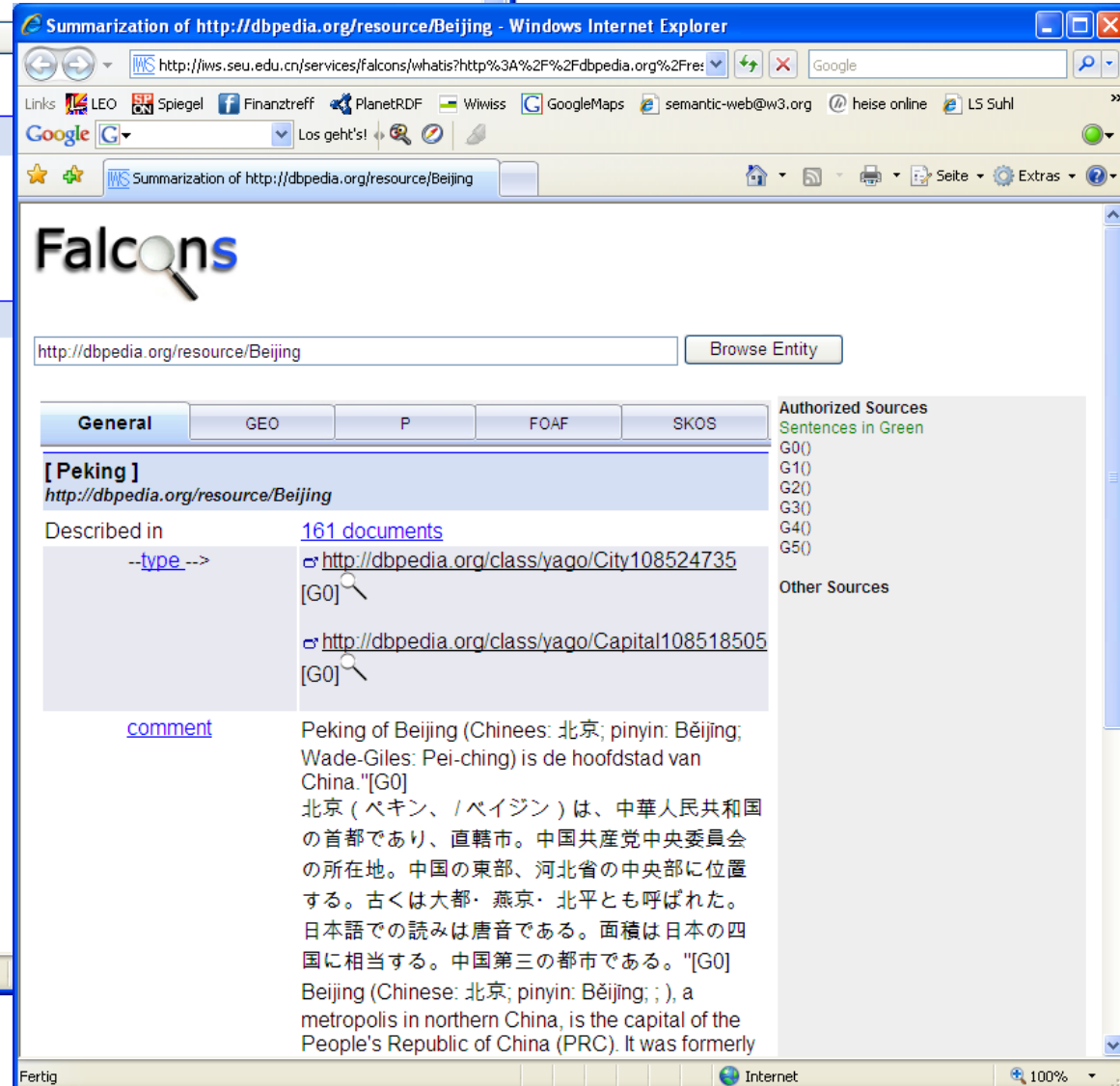
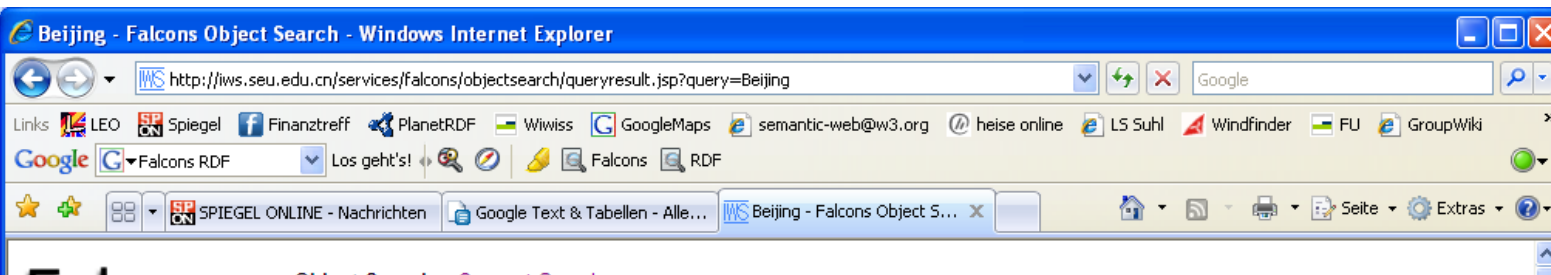
- Geospatial entry point into the Web of Data
- Uses DBpedia, Revyu and Flickr



Web of Data Search Engines

- SWSE (DERI, Ireland)
- Swoogle (UMBC, USA)
- **Falcons (IWS, China)**
- **Sindice (DERI, Ireland)**
- **Watson (Open University, UK)**
- **MicroSearch (Yahoo, Spain)**

Falcons



Search Results 🍌

Persons: 3 vCards: 35 Events: 25 Unfinished: 0 | 1 - 10 of about 2244

1. [Ivan's private site](#)

<http://www.ivan-herman.net/> - 46k [Update metadata](#)

2. [Ivan Herman](#)



name: Ivan Herman

personal mailbox: <mailto:ivan@w3.org>

homepage: <http://www.ivan-herman.net>

Ivan Herman gives a talk on behalf of the China Office entitled "What is the ... Ben Adida, Elias Torres, and Ivan Herman give a tutorial entitled "RDFa: ...

<http://www.w3.org/People/ivan/> - 25k [Update metadata](#)

[homepage](#)

[Ivan's private site](#)

<http://www.ivan-herman.net/> - 46k [Update metadata](#)

3. [Amsterdam, the Netherlands \(the city where I live...\)](#)

Lots of pictures taken when walking around this very ... Total images: 79 | Last update: 2008-02-16 | © JAlbum & Chameleon | Help © Ivan Herman, 2007 ...

<http://www.ivan-herman.net/Photos/Amsterdam/> - 53k [Update metadata](#)

4. [Ivan Herman](#)



3. What is next?

Publish more datasets!

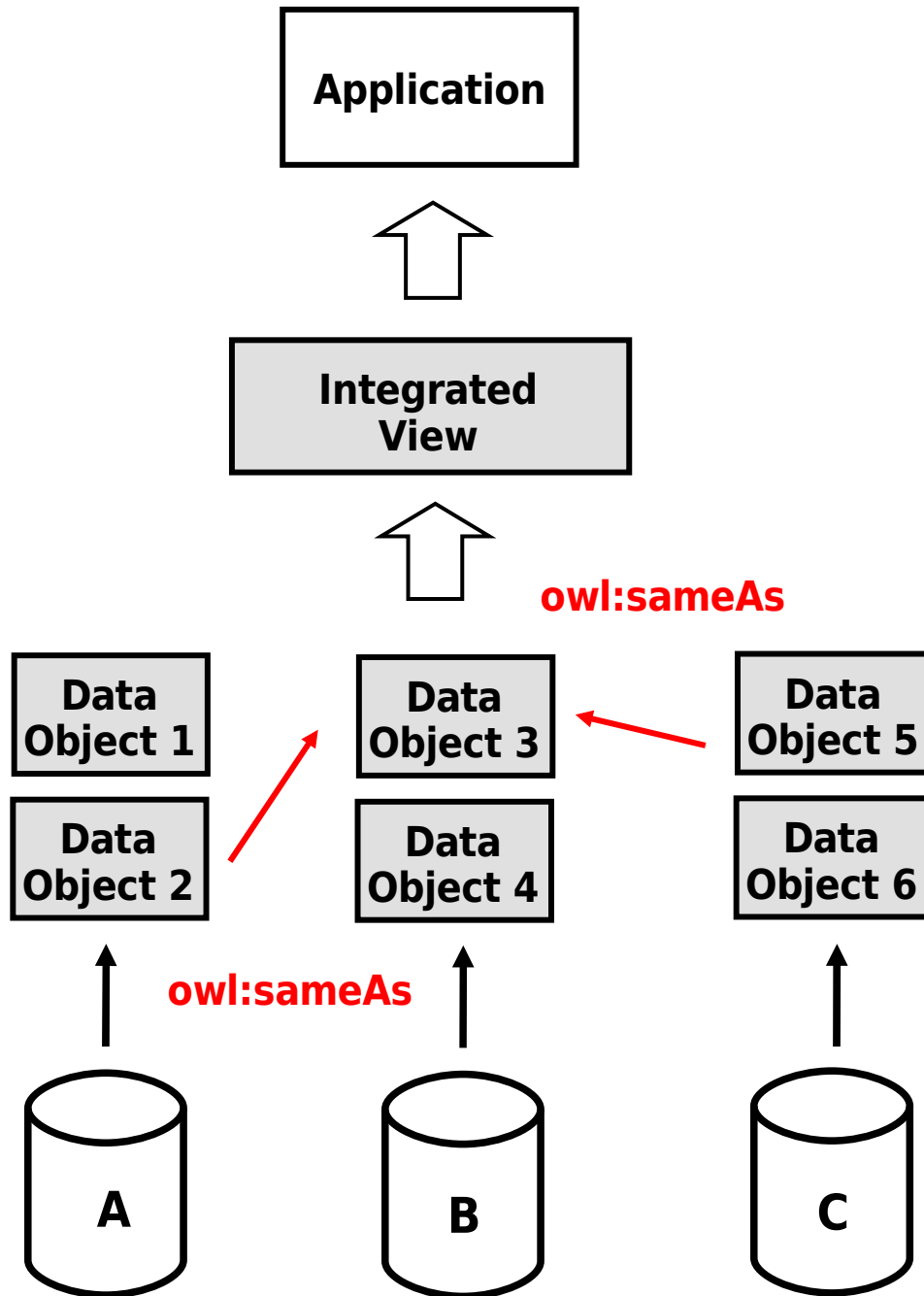
- 1. Conversion of further open license datasets into RDF**
- 2. Wrappers around existing applications**

- **Tutorial: How to publish Linked Data on the Web**
- **LOD Triplification Challenge at I-Semantics 2008**
 - **Win a MacBook Air, Asus EeePC, iPod Touch**
 - **Deadline: June 30th, 2008**

Linking

- 1. Increase the amount of links between datasets**
 - 2. Increase the quality of these links**
- Today: Simple pattern- and graph-matching based techniques used for automated interlinking.**
 - There is lots of existing work in database and knowledge representation communities on identity resolution to be used.**

Data Fusion



Users want an integrated view on all data that is available about an object!

- **Raises well known but unsolved problems:**
 - Schema mapping
 - Inconsistency resolution
 - Trust / information quality

In order to do anything serious with data from the Web, its license terms have to be clear.

■ Need for

- **proper licensing vocabularies for dedicating data to the public domain**
- **best practices on how to annotate data with licensing meta-data**

■ Can build on

- **Open Data Commons Public Domain Dedication & Licence (PDDL)**
- **Creative Commons Licensing Framework**

Browsers and Search Engines FOR THE END USER

We need real tools, not only proof of concept prototypes!

- **End user friendly views on the data**
 - ordering and merging of properties
 - dealing with information overflow
- **More advanced data analysis features**
 - aggregation, drill down
 - calculations, Web-Excel
- **Explanations about data provenance and trustworthiness**

Participating in the Linking Open Data Project

■ Wiki Page

- <http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>

■ Mailing List

- public-lod@w3.org
- <http://lists.w3.org/Archives/Public/public-lod/>

■ Participating in the project

- Put your name on the Wiki page
- Subscribe to the mailing list
- Do something useful

Thanks!

References

- **Design Issues: Linked Data**
<http://www.w3.org/DesignIssues/LinkedData.html>
- **Tutorial on How to Publish Linked Data on the Web**
<http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/>
- **Linking Open Data Project Wiki**
<http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>