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



Overview

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?

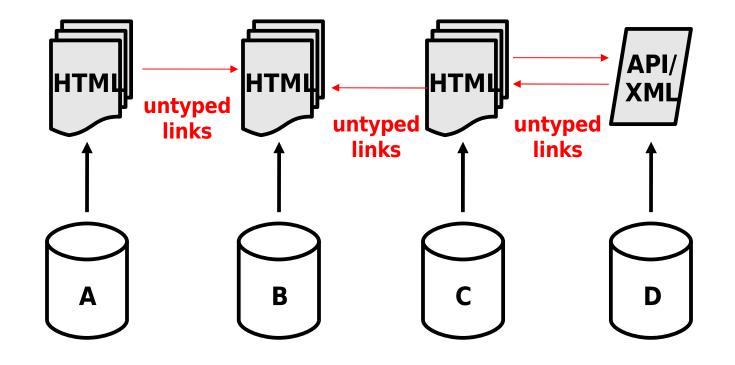
Open Issues and directions for future work

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



Image: Bob Jagensdorf, http://flickr.com/photos/darwinbell/, CC-BY

How do you know these documents are about Beijing?

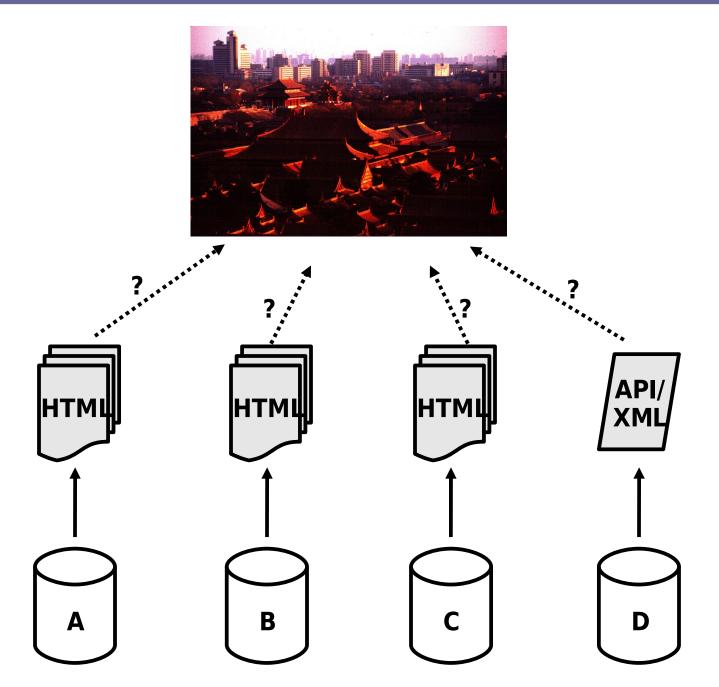


Image: Paul Downey, http://flickr.com/photos/psd/, CC-BY

Disconnected Data



Data Integration

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

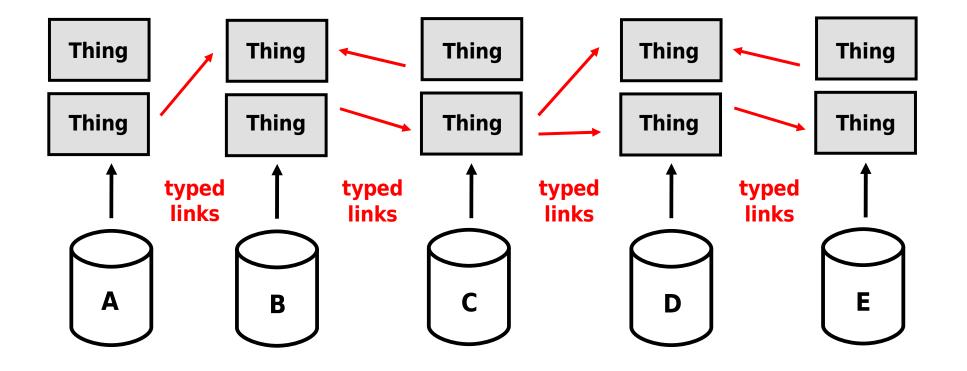
Querying Across Data Sources

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

Many common things are represented in multiple data sets

Linking identifiers connects these data sets

Linked data opens the doors of the silos



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



Image: Steve Jurvetson, http://www.flickr.com/people/jurvetson/, CC-BY Christian Bizer, Tom Heath, Tim Berners-Lee: Linking Open Data (2/19/2008)

- **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

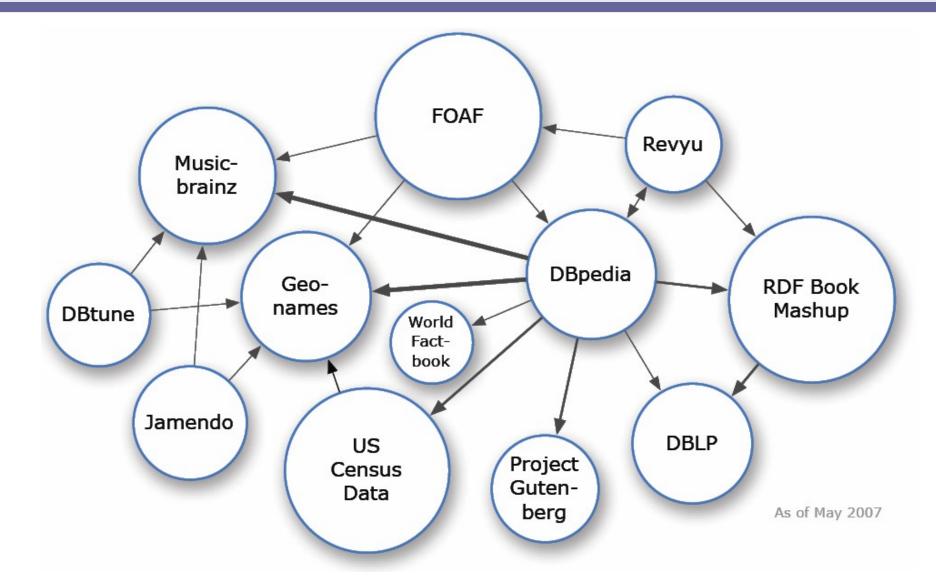
2. W3C SWEO Linking Open Data Project



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

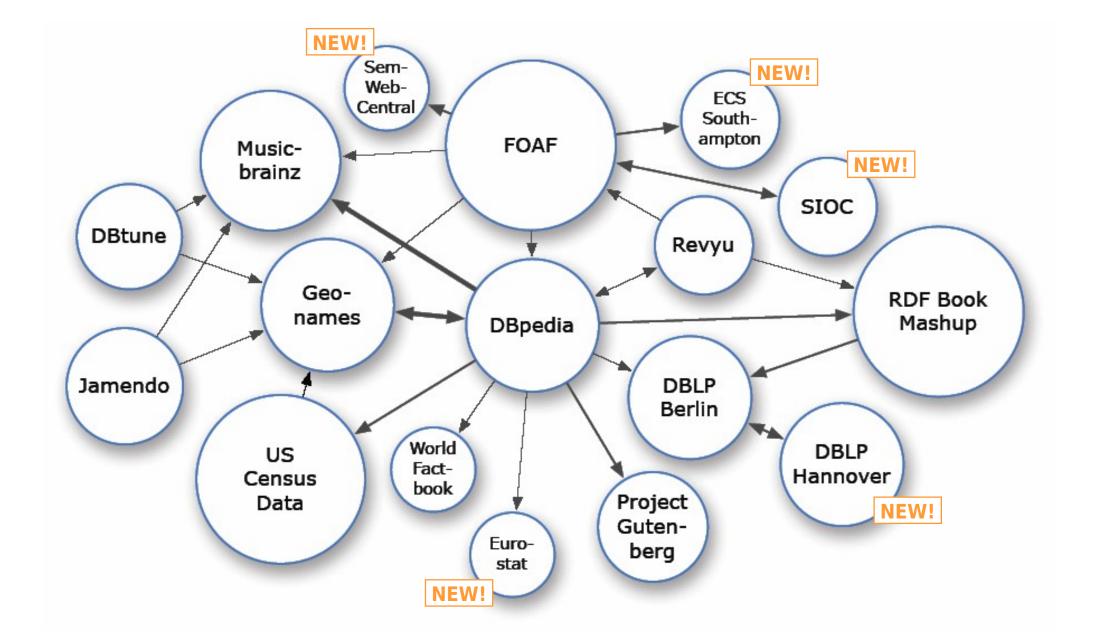
LOD Datasets on the Web: May 2007



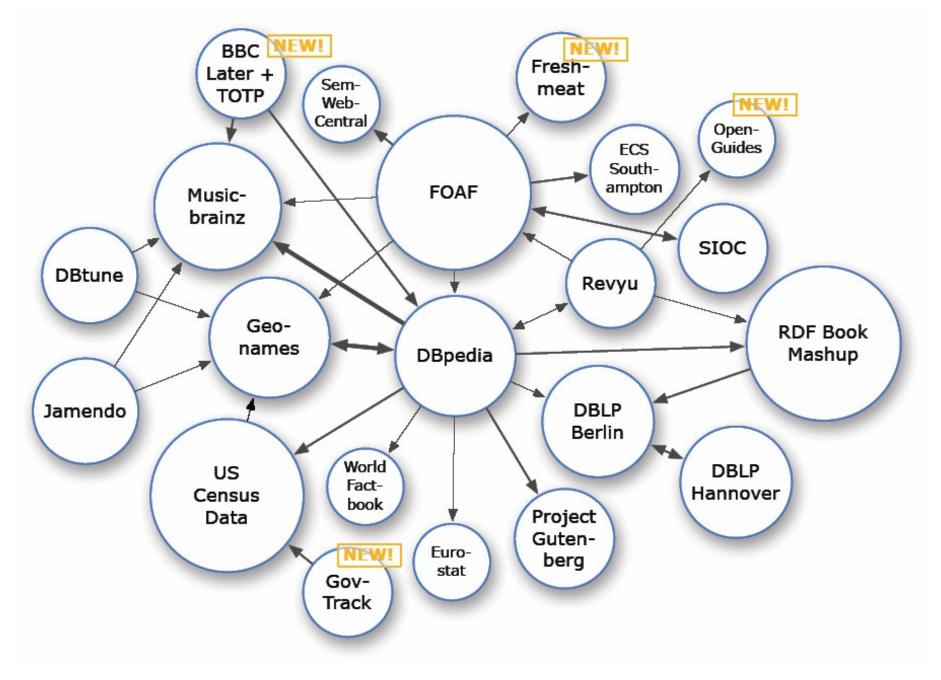
Over 500 million RDF triples

Around 120,000 RDF links between data sources

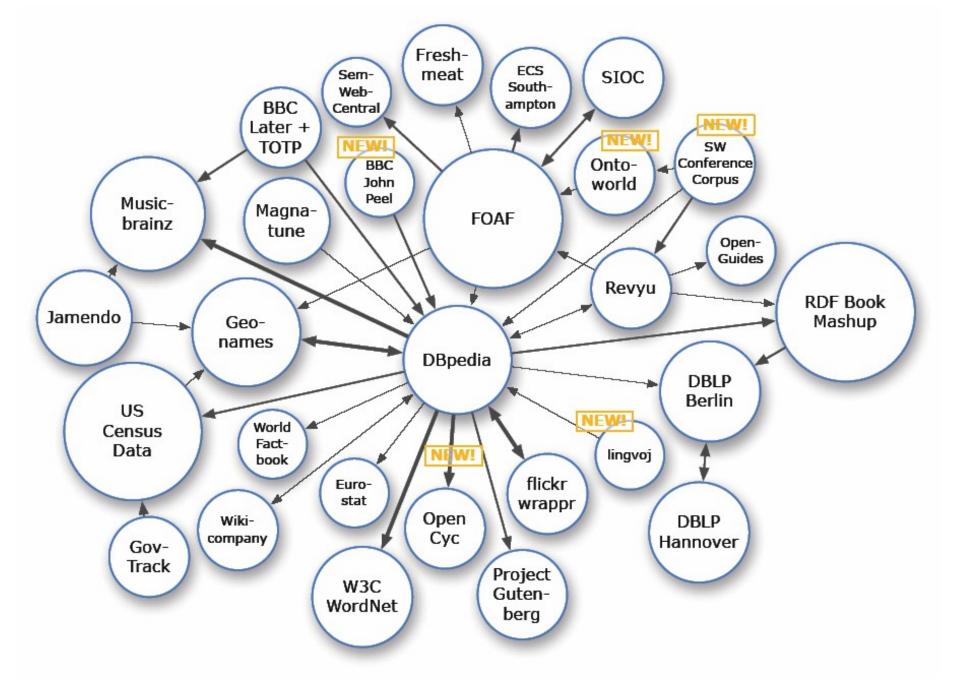
LOD Datasets on the Web: July 2007



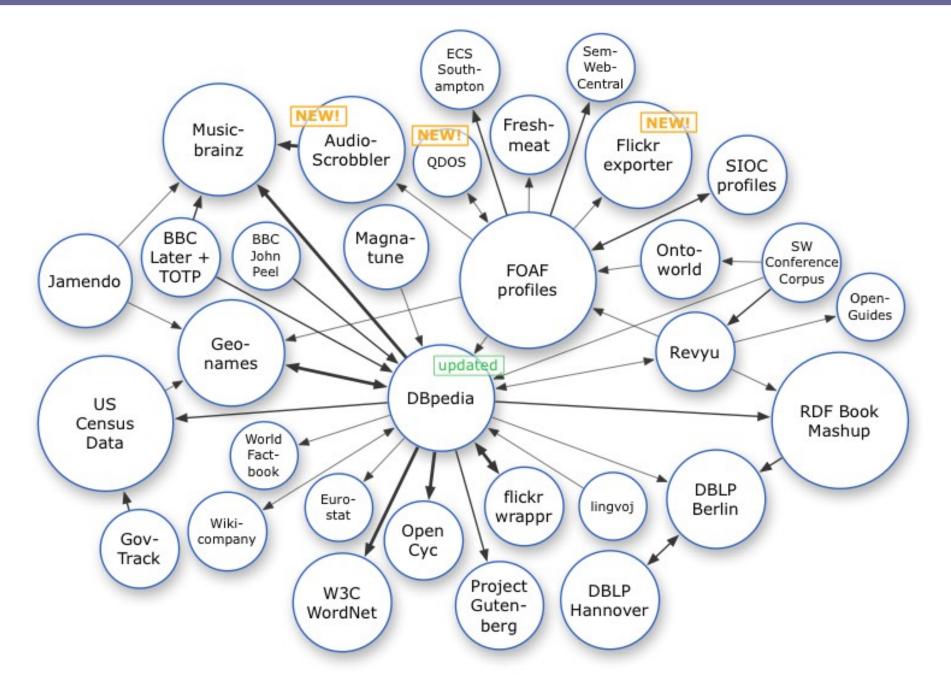
LOD Datasets on the Web: August 2007



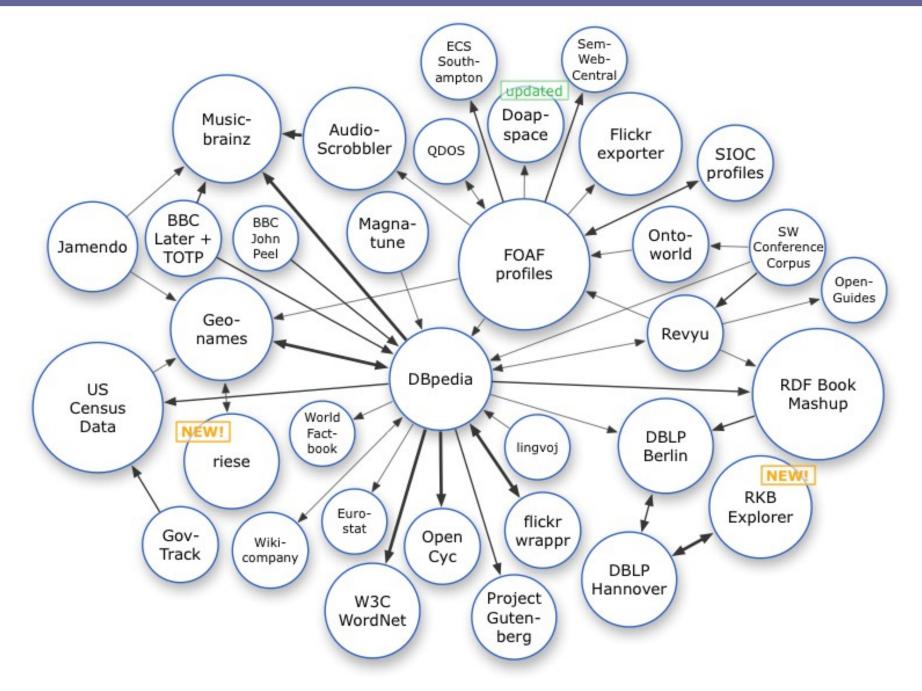
LOD Datasets on the Web: November 2007



LOD Datasets on the Web: February 2008



LOD Datasets on the Web: April 2008



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

(rough estimates)

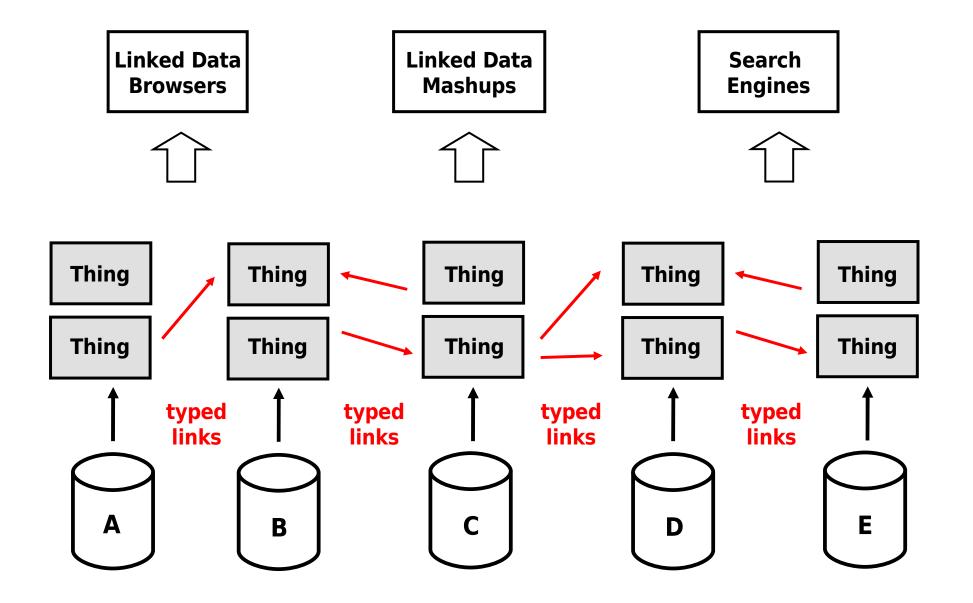
Organizations participating in the LOD community

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?



- 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, Irland)
- Marbles (FU Berlin, DE)

Open



http://www.w3.org/1999/02/22-rdf-syntax-ns#type	Building 💥
	• Feature
comment	 Der Flughafen Peking ist einer der wichtigsten Flughafen 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 distict de
s <u>sameAs</u> of	Beijing Capital International Airport
<u>depiction</u>	
	CHINA MOBILE

:

URI: http://www.w3.org/People/Berners-Lee/card#i

Add to Outliner Help About

This is release 0.8 of the Tabulator Project. The live development trunk is also available.

▼Tim Berners-Lee Request for <http://www4.wiwiss.fu-berlin.de/bookmashup/persons/Tim+Berners-Lee> request Request for <http://www4.wiwiss.fu-berlin.de/dblp/resource/person/100007>. requested by ▶ ● Tim Berners-Lee's FOAF file 🖹 ▶ OMale type ▶ ● Person 🗐 Tim Berners-Lee label seeAlso ORDF document describing: Tim Berners-Lee ▶ ● RDF Description of Tim Berners-Lee OAmy van der Hiel assistant homePage Ohttp://www.w3.org/People/Berners-Lee/ work ▶... organization World Wide Web Consortium title W3C (large: includes all specs and groups) SocialEntity type ▶ Corganization 🗐 W3C label seeAlso ▶●W3C Groups and Organizational Structure 🖹 OVV3C Standards and Technical Reports publicHomePage > Ohttp://www.w3.org/ ▶ Shttp://www.w3.org/ 🗐 homepage logo World Wide Web Consortium name is organization of ▶ © Tim Berners-Lee publicHomePage ▶ Ohttp://www.w3.org/People/Berners-Lee/ 🗐

Domain-specific applications using Linked Data from the Web

Revyu

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

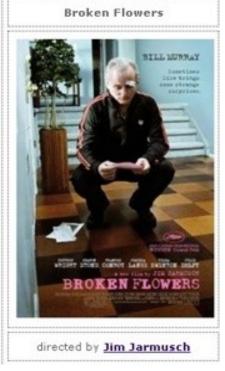
bill-murray film jessica-lange jim-jarmusch julie-delpy movie sharon-stone

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 ...



DBpedia Mobile



Geospatial entry point into the Web of Data

Uses DBpedia, Revyu and Flickr



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

Falcons

🖉 Beijing - Falcons Object Search - Windows Internet Explorer					
🚱 🕞 👻 🔣 http://iws.seu.edu.cn/services/falcons/objectsearch/queryresult.jsp?query=Beijing					
Links K LEO R Spiegel F Finanztreff and PlanetRDF - Wiwiss C Google Maps & semantic-web@w3.org @ heise online & LS Suhl A Windfinder - FU & GroupWiki * Google C + Falcons RDF v Los geht's! • & Ø Ø E Falcons RDF v Los geht's! • Ø Ø Ø E Falcons RDF v Los geht's! • Ø Ø Ø E Falcons RDF v Los geht's! • Ø Ø Ø E Falcons RDF v Los geht's! • Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø					
🚖 🏟 🔠 🗸 🚼 🛪 Ele 🗸 🔂 😪 ele z 🚱 Seite z 🎯 Extras z 🔞 z					
		<u> </u>			
Faicons Object Search Concept Search Search Objects	C Summarization of http://dl	bpedia.org/resource/Beijing - Windows Internet Explorer			
Supports Boolean operators, quotes, and wildcard characters.	💽 🗸 🔣 http://iws.seu.ed	du.cn/services/falcons/whatis?http%3A%2F%2Fdbpedia.org%2Fre: 💌 🛃			
	Links 🌿 LEO 🔣 Spiegel 👔 Finanztreff 🐗 PlanetRDF 🖃 Wiwiss 🕞 GoogleMaps 🖉 semantic-web@w3.org 🖉 heise online 🍃 LS Suhl 🛛 🎇				
All	Google G-	🔽 Los geht's! 🗄 🍭 🖉 🦼	<u>_</u>		
Artifact Capital City Document Group	😪 🏟 🔣 Summarization of http	p://dbpedia.org/resource/Beijing	• 🔊 🔹 🖶 • 🔂 Seite • 🍈 Extras • 🕢		
Institution Landmark Location Noun Synset Ontology Organization Person Publication Subject System	Falcons		~		
Objects 1 - 10 of 8634 for your search Beijing (1.223 seconds)					
	http://dbpedia.org/resource/Beijing Browse Entity				
Beijing Types: Capital, City	General GE	O P FOAF SKOS	Authorized Sources Sentences in Green		
L abels : 北京" Pekin Пекин" 北京市" Pequim Pechino Beijing Pékin" Peking Pekín" http://dbpedia.org/resource/ Beijing - <u>Described in 184 documents</u>	[Peking] G00 http://dbpedia.org/resource/Beijing G20 G30 G30				
Beijing	Described in	161 documents	G4() G5()		
Types: Subject, Labels: Beijing	<u>type</u> >		Other Sources		
http://ontoworld.org/wiki/Special:URIResolver/Beijing - Described in 11 documents		a http://dbpedia.org/class/yago/Capital108518505			
Beijing Guoan		[G0]			
Types: Club Labels: Beijing Hyundai 北京国安" 北京国安足球俱乐部" Beijing Guoan	<u>comment</u>	Peking of Beijing (Chinees: 北京; pinyin: Běijīng;			
http://dbpedia.org/resource/Beijing_Guoan - Described in 30 documents		Wade-Giles: Pei-ching) is de hoofdstad van China."[G0] 北京(ペキン、/ ペイジン)は、中華人民共和国			
Beijing		の首都であり、直轄市。中国共産党中央委員会			
Types: Capital City		の所在地。中国の東部、河北省の中央部に位置			
Labels: Beijing http://lonely.org/russia#Beijing - <u>Described in 5 documents</u>		する。古くは大都・燕京・北平とも呼ばれた。 日本語での読みは唐音である。面積は日本の四			
		日本品での読みは居自てのる。画復は日本の日 国に相当する。中国第三の都市である。"[G0]			
ertig		Beijing (Chinese: 北京; pinyin: Běijīng; ;), a			
		metropolis in northern China, is the capital of the People's Republic of China (PRC). It was formerly			
	Fertig	ee Inte	rnet 🔍 100% 🔻 🛒		



Search

E

Search Results 🗳

Ivan's private site 1

http://www.ivan-herman.net/ - 46k Update metadata

Ivan Herman 2



ame: 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/lvan/ - 25k Update metadata

homepage

Ivan's private site

http://www.ivan-herman.net/ - 46k Update metadata

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

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



Ivan Herman

3. What is next?

Conversion of further open license datasets into RDF Wrappers around existing applications

Tutorial: How to publish Linked Data on the Web

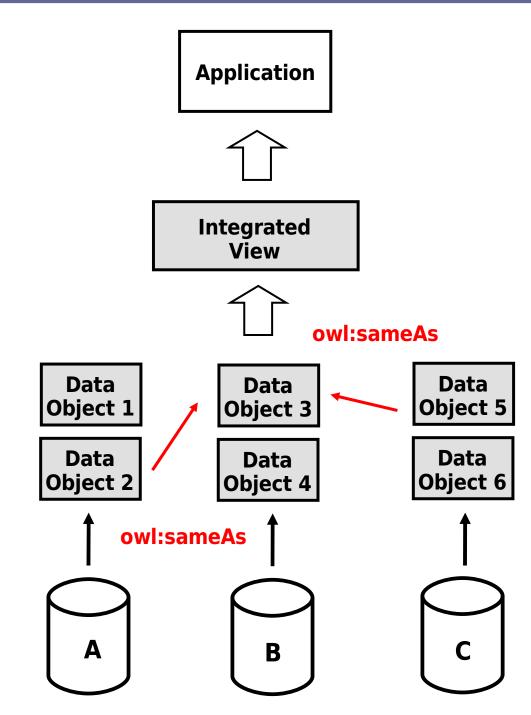
LOD Triplification Challenge at I-Semantics 2008

- Win a MacBook Air, Asus EeePC, iPod Touch
- Deadline: June 30th, 2008



- Increase the amount of links between datasets
 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 metadata

Can build on

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

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

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