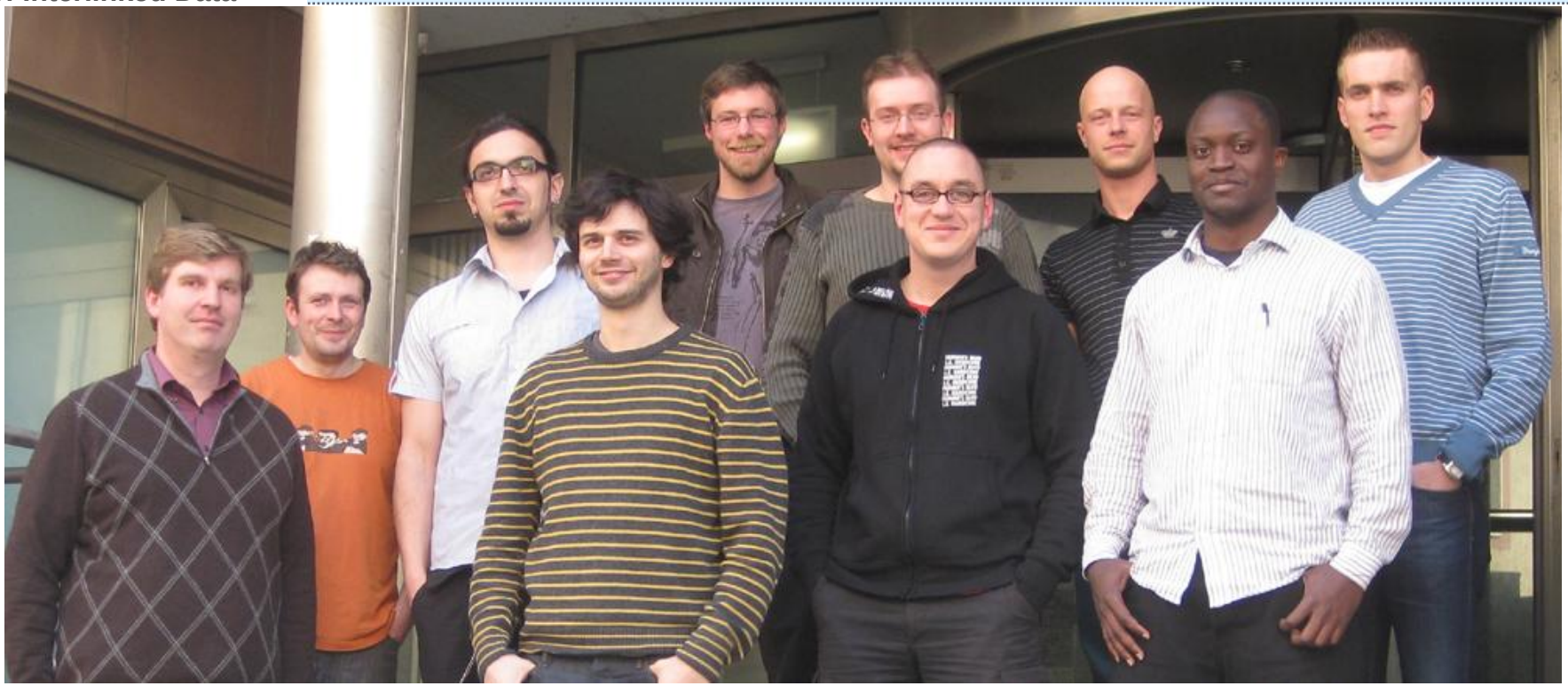
The Linked Data Lifecycle

Jens Lehmann, Sören Auer

UNIVERSITÄT LEIPZIG

LOD2
Creating Knowledge
out of Interlinked Data

AKSW Group



- 3 PostDocs, 15 PhDs, 3 Subgroups (Emergant Semantics, MOLE,SIMBA)
- centered around Semantic Web tools and applications
- http://aksw.org
- http://www.slideshare.net/soeren1611/overview-ag-aksw

AKSW Projects

Active Funded Projects:

- LOD2
- LATC
- SCMS
- LE4SW

Research/Software Projects:

- DBpedia
- LinkedGeoData
- OntoWiki
- DL-Learner
- BorderFlow
- LIMES
- SAIM
- NLP2RDF
- Triplify
- . . .

Personal Background

- Studied Computer Science at TU Dresden 2001-2006
- PhD Uni Leipzig 2006-2010
- PostDoc AKSW / Uni Leipzig ab 2010
- Head of Machine Learning and Ontology Engineering Group (MOLE)
- EU-Projects: LOD2, LATC
- Software/Research-Projects: DBpedia, DL-Learner, ORE, LinkedGeoData
- Interests: Semantic Web, Machine Learning, Logic

Why Semantic Web?

Why Semantic Web?

Problem: Try to search for these things on the current Web:

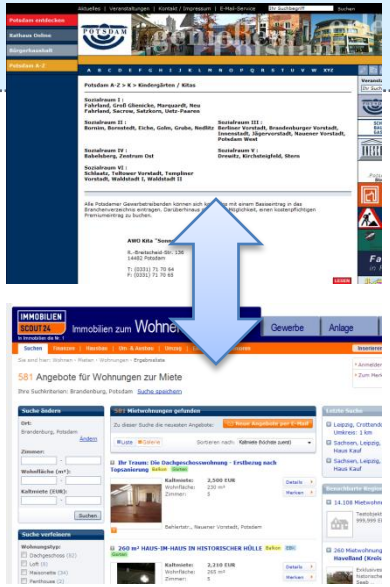
- Apartments near German-Russian bilingual childcare in Leipzig.

Why Semantic Web?

Problem: Try to search for these things on the current Web:

- Apartments near German-Russian bilingual childcare in Leipzig.
- ERP service providers with offices in Vienna and London.

Why Semantic Web?

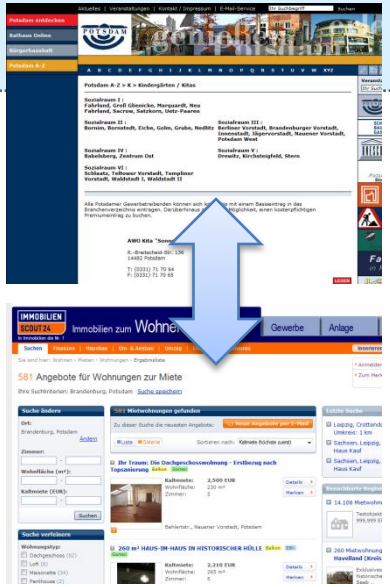


Problem: Try to search for these things on the current Web:

- Apartments near German-Russian bilingual childcare in Leipzig.
- ERP service providers with offices in Vienna and London.
- Researchers working on multimedia topics in Eastern Europe.

Information is available on the Web, **but opaque to current Web search**.

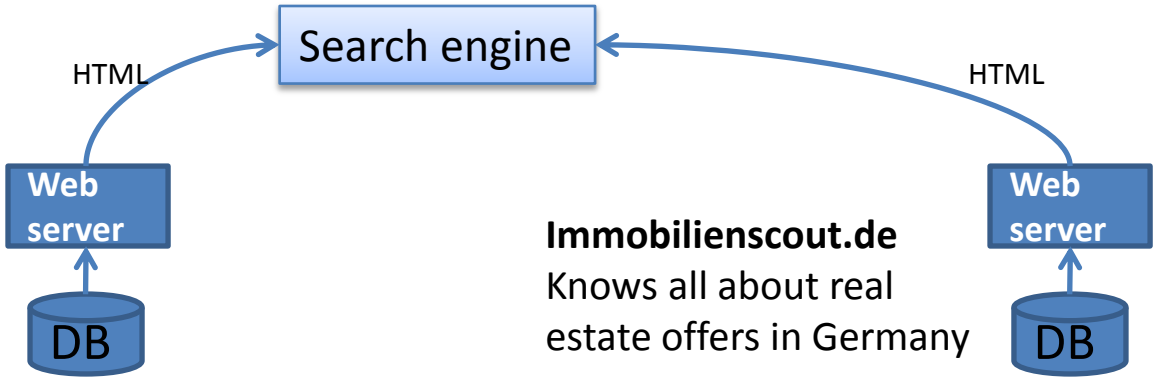
Why Semantic Web?



Problem: Try to search for these things on the current Web:

- Apartments near German-Russian bilingual childcare in Leipzig.
- ERP service providers with offices in Vienna and London.
- Researchers working on multimedia topics in Eastern Europe.

Information is available on the Web, **but opaque to current Web search**.

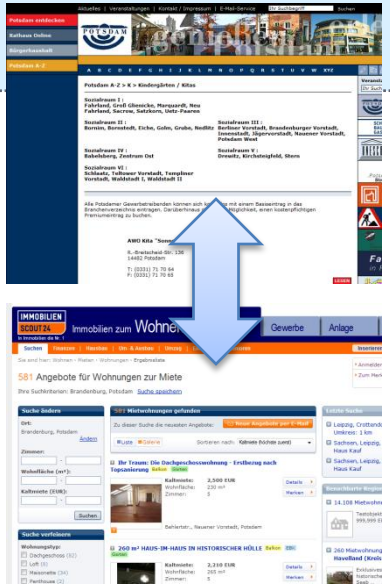


leipzig.de
Has everything about childcare in Leipzig.

Immobilienscout.de
Knows all about real estate offers in Germany

LOD2
Creating Knowledge
out of Interlinked Data

Why Semantic Web?

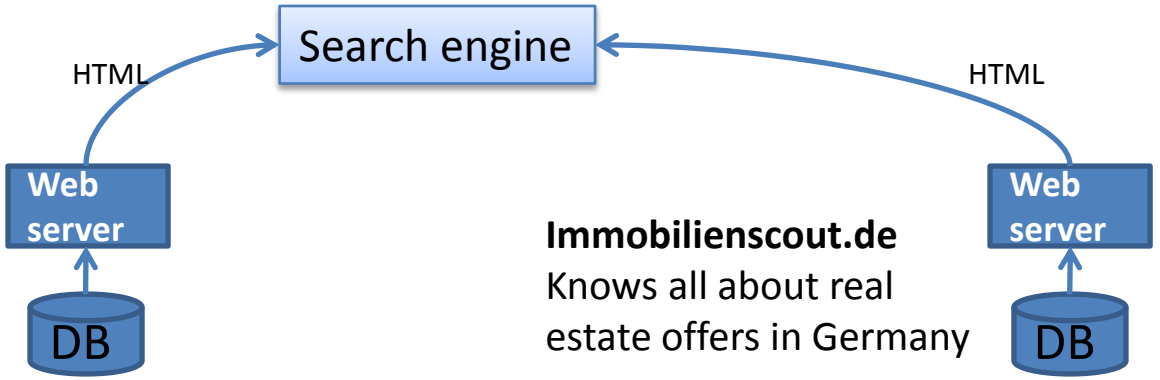


Problem: Try to search for these things on the current Web:

- Apartments near German-Russian bilingual childcare in Leipzig.
- ERP service providers with offices in Vienna and London.
- Researchers working on multimedia topics in Eastern Europe.

Information is available on the Web, **but opaque to current Web search**.

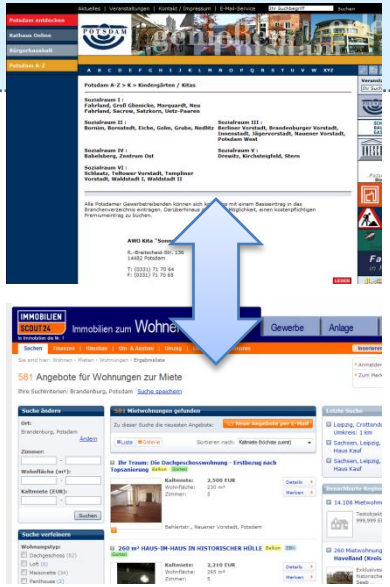
Solution: complement text on Web pages with structured linked open data & intelligently combine/integrate such structured information from different sources:



leipzig.de
Has everything about childcare in Leipzig.

Immobilienscout.de
Knows all about real estate offers in Germany

Why Semantic Web?

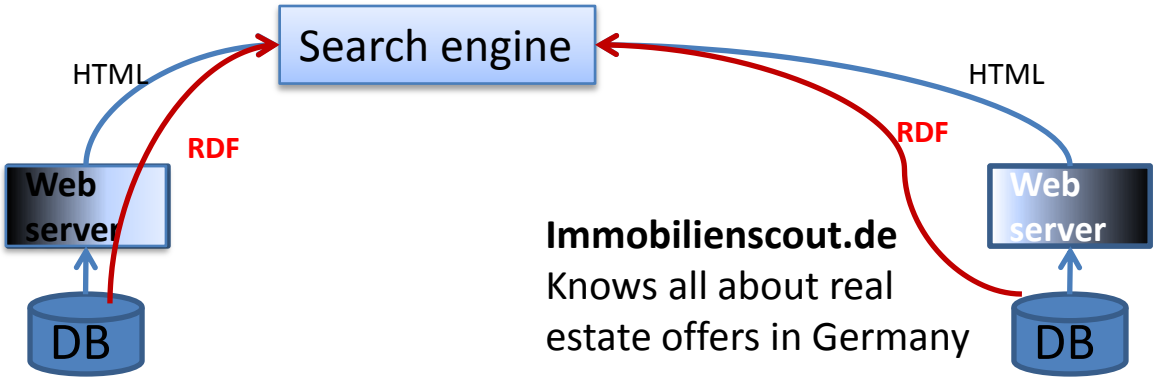


Problem: Try to search for these things on the current Web:

- Apartments near German-Russian bilingual childcare in Leipzig.
- ERP service providers with offices in Vienna and London.
- Researchers working on multimedia topics in Eastern Europe.

Information is available on the Web, **but opaque to current Web search**.

Solution: complement text on Web pages with structured linked open data & intelligently combine/integrate such structured information from different sources:



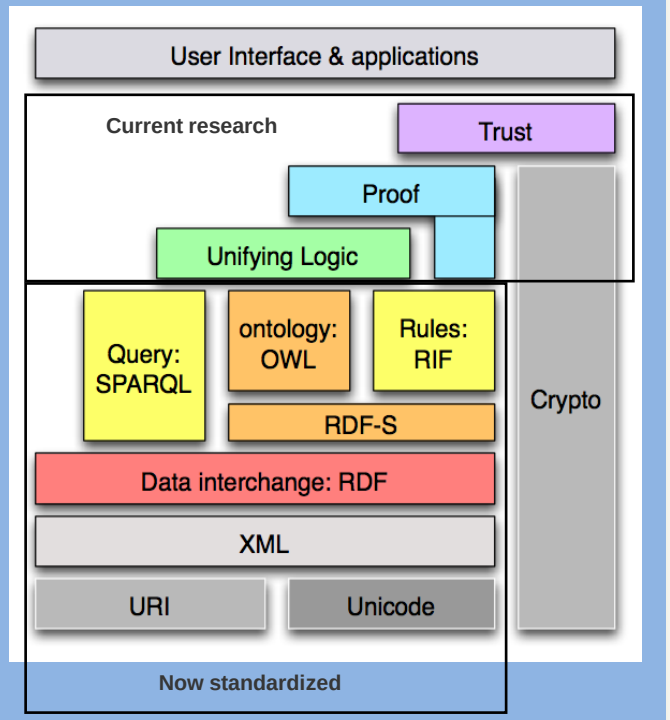
leipzig.de
Has everything about childcare in Leipzig.

Immobilienscout.de
Knows all about real estate offers in Germany

LOD2
Creating Knowledge
out of Interlinked Data

Semantic Web - Standards

Semantic Web Architecture



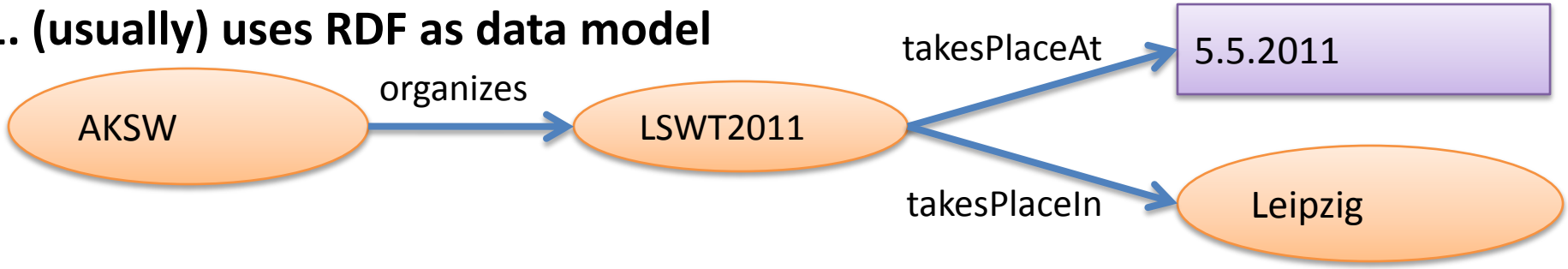
Standardization Semantic Web

1994	• First public presentation of idea
1998	• Start of standardization of RDF and RDFS
2000	• large research projects in US and Europe
2002	• Start of OWL standardisation
2004	• RDF • OWL
2006	
2008	•RDFa •SPARQL
2009	• OWL 2
2010	•RIF

W3C Semantic Web

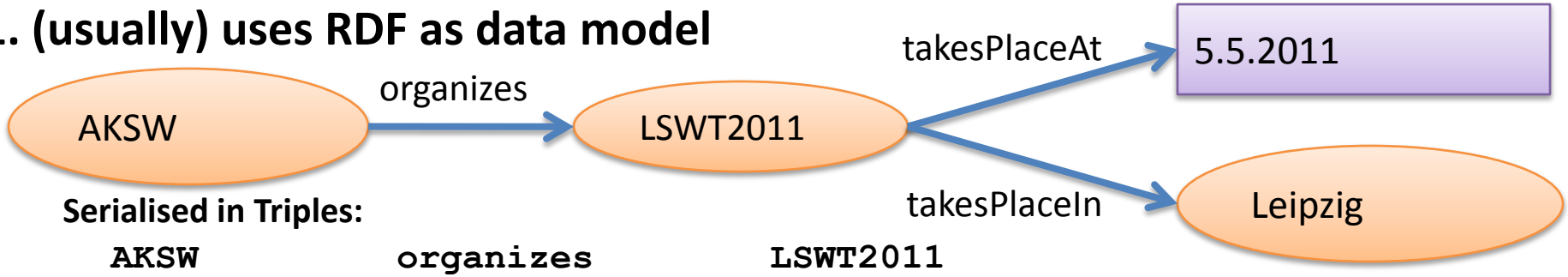
Linked Data

1. (usually) uses RDF as data model



Linked Data

1. (usually) uses RDF as data model

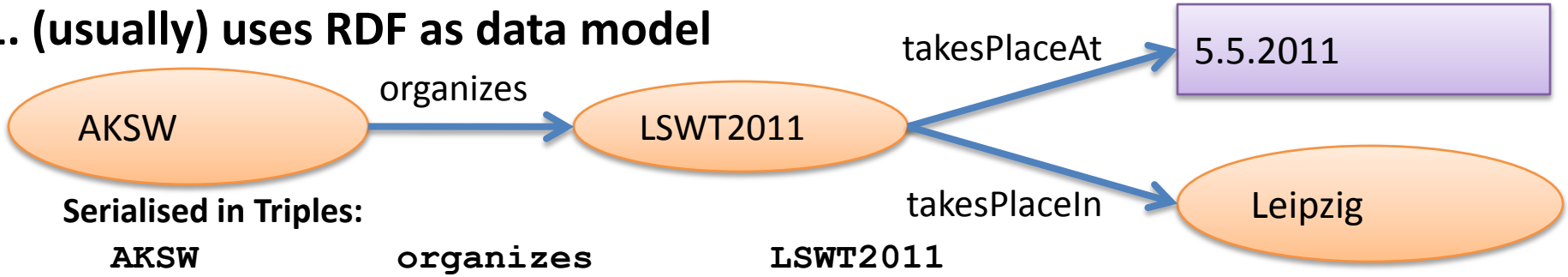


Serialised in Triples:

```
AKSW        organizes       LSWT2011
LSWT2011    takesPlaceAt    "20110505"^^xsd:date
LSWT2011    takesPlaceIn    Leipzig
```

Linked Data

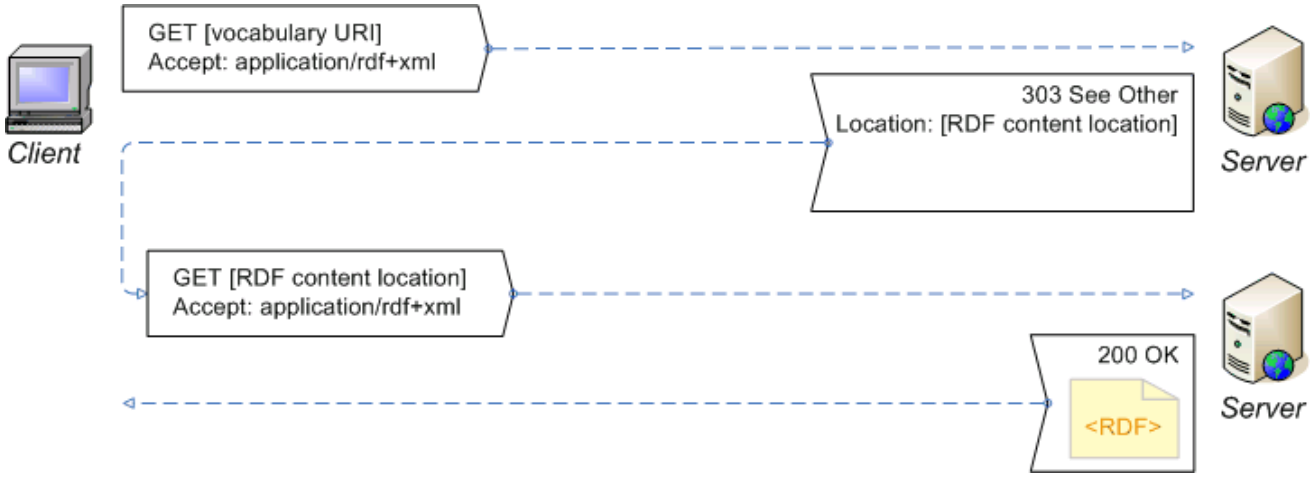
1. (usually) uses RDF as data model



Serialised in Triples:

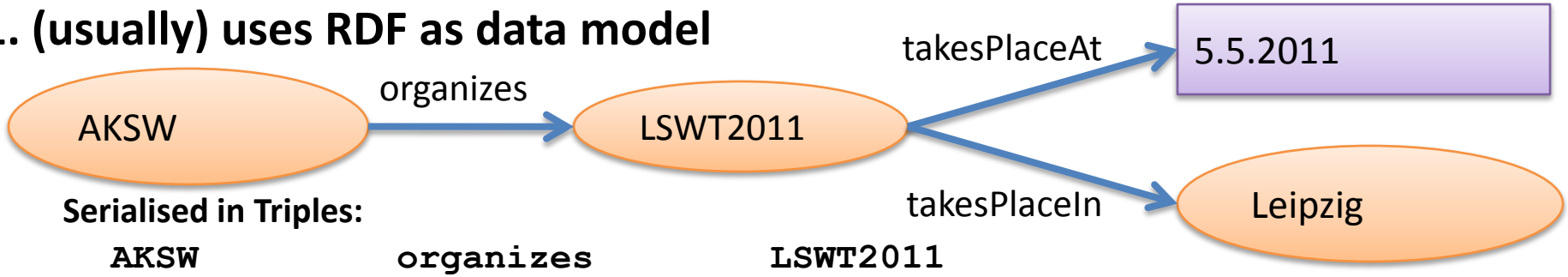
```
AKSW          organizes       LSWT2011
LSWT2011      takesPlaceAt    "20110505"^^xsd:date
LSWT2011      takesPlaceIn    Leipzig
```

2. Uses dereferencable HTTP-URIs, e.g. Content-negotiation



Linked Data

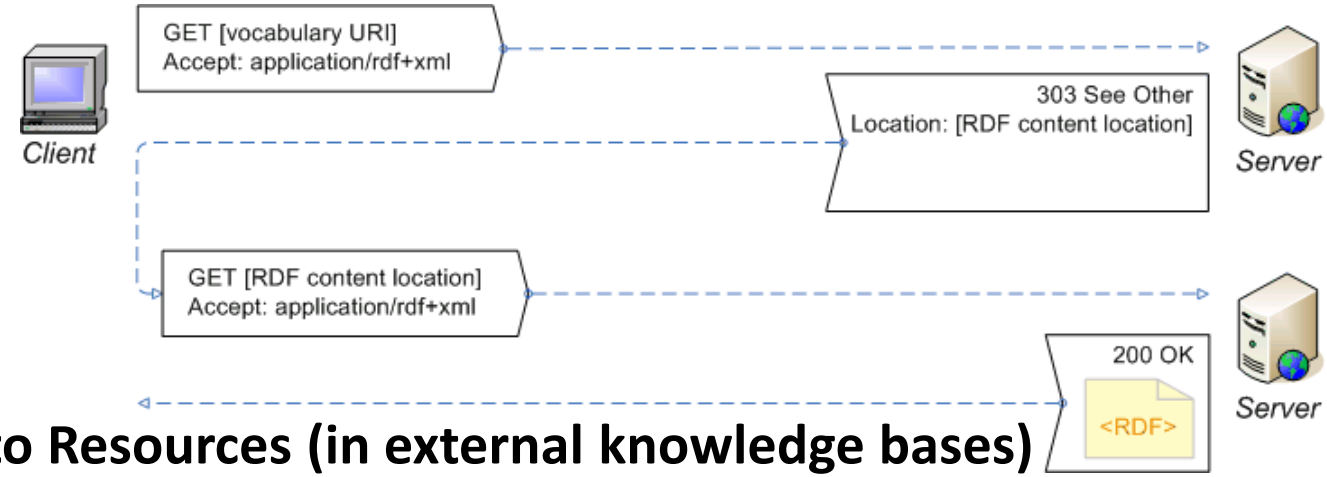
1. (usually) uses RDF as data model



Serialised in Triples:

```
AKSW        organizes       LSWT2011
LSWT2011    takesPlaceAt    "20110505"^^xsd:date
LSWT2011    takesPlaceIn    Leipzig
```

2. Uses dereferencable HTTP-URIs, e.g. Content-negotiation

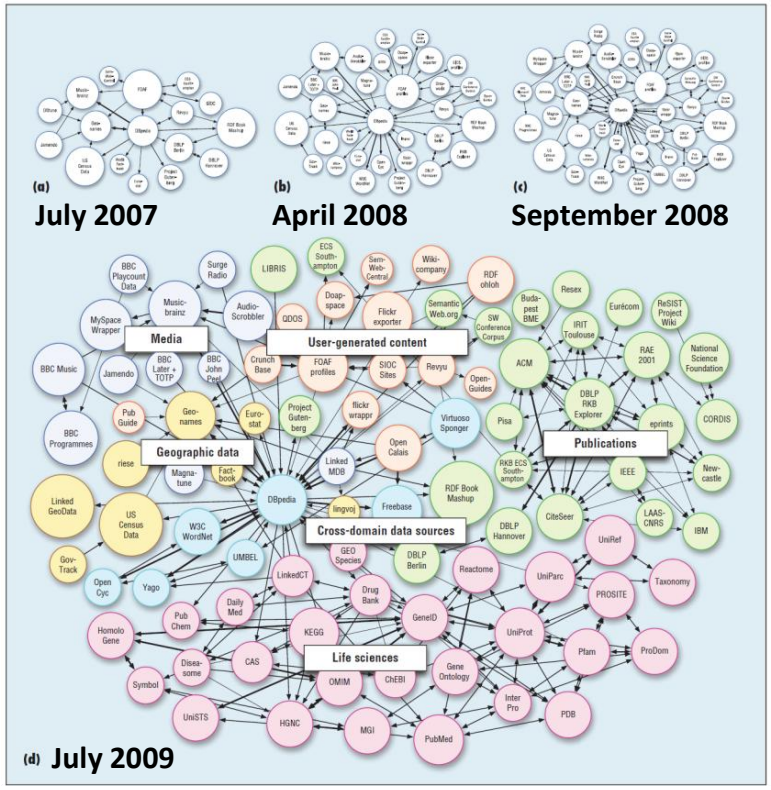


3. Links to Resources (in external knowledge bases)

LOD2
Creating Knowledge
out of Interlinked Data

LOD achievements and challenges

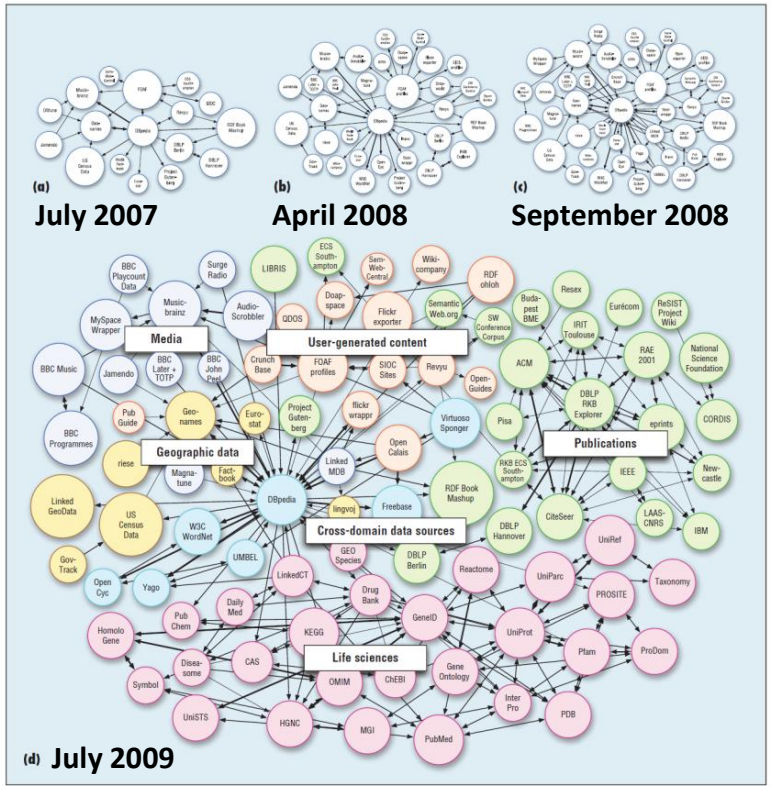
- Web - a global, distributed platform for data, information and knowledge integration
- exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF



LOD2
Creating Knowledge
out of Interlinked Data

LOD achievements and challenges

- Web - a global, distributed platform for data, information and knowledge integration
- exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF



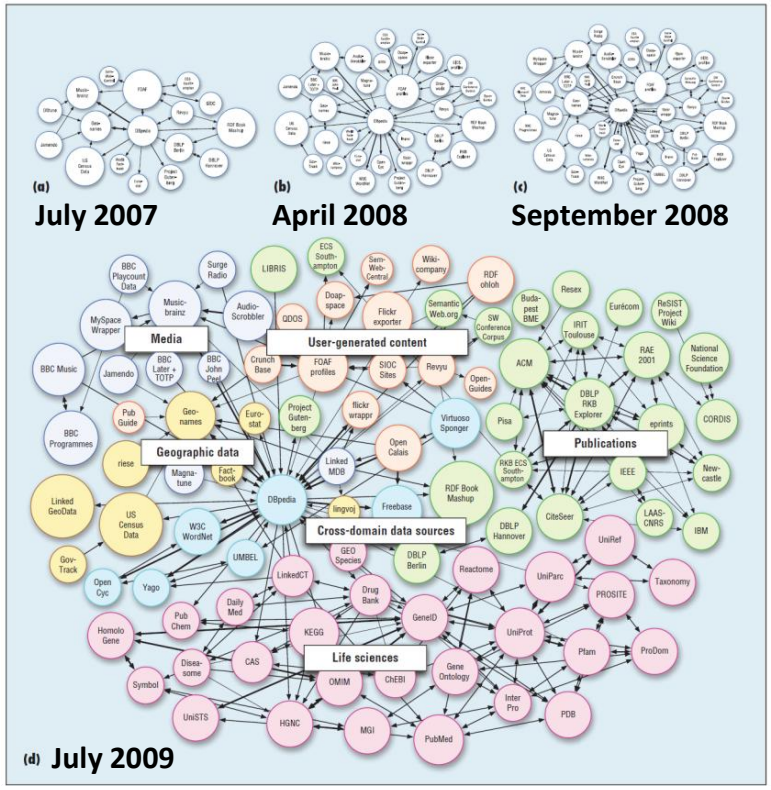
Achievements

1. Extension of the Web with a **data commons** (25B facts)
2. vibrant, global RTD **community**
3. **Industrial uptake** begins (e.g. BBC, Thomson Reuters, Eli Lilly, NY Times, Facebook, Google, Yahoo)
4. Emerging **governmental adoption** in sight
5. Establishing Linked Data as **a deployment path** for the Semantic Web.

LOD2
Creating Knowledge out of Interlinked Data

LOD achievements and challenges

- Web - a global, distributed platform for data, information and knowledge integration
- exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF



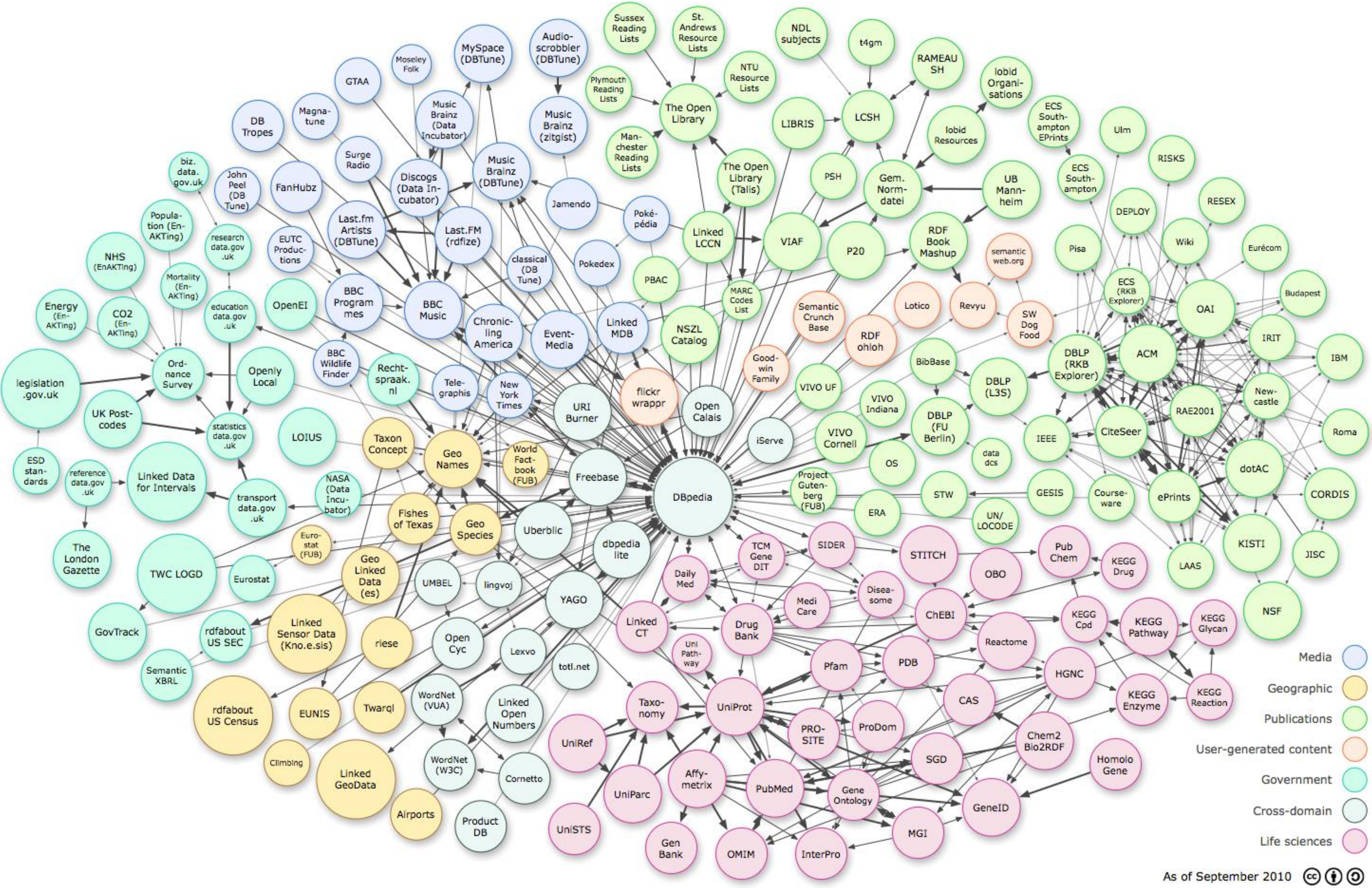
Achievements

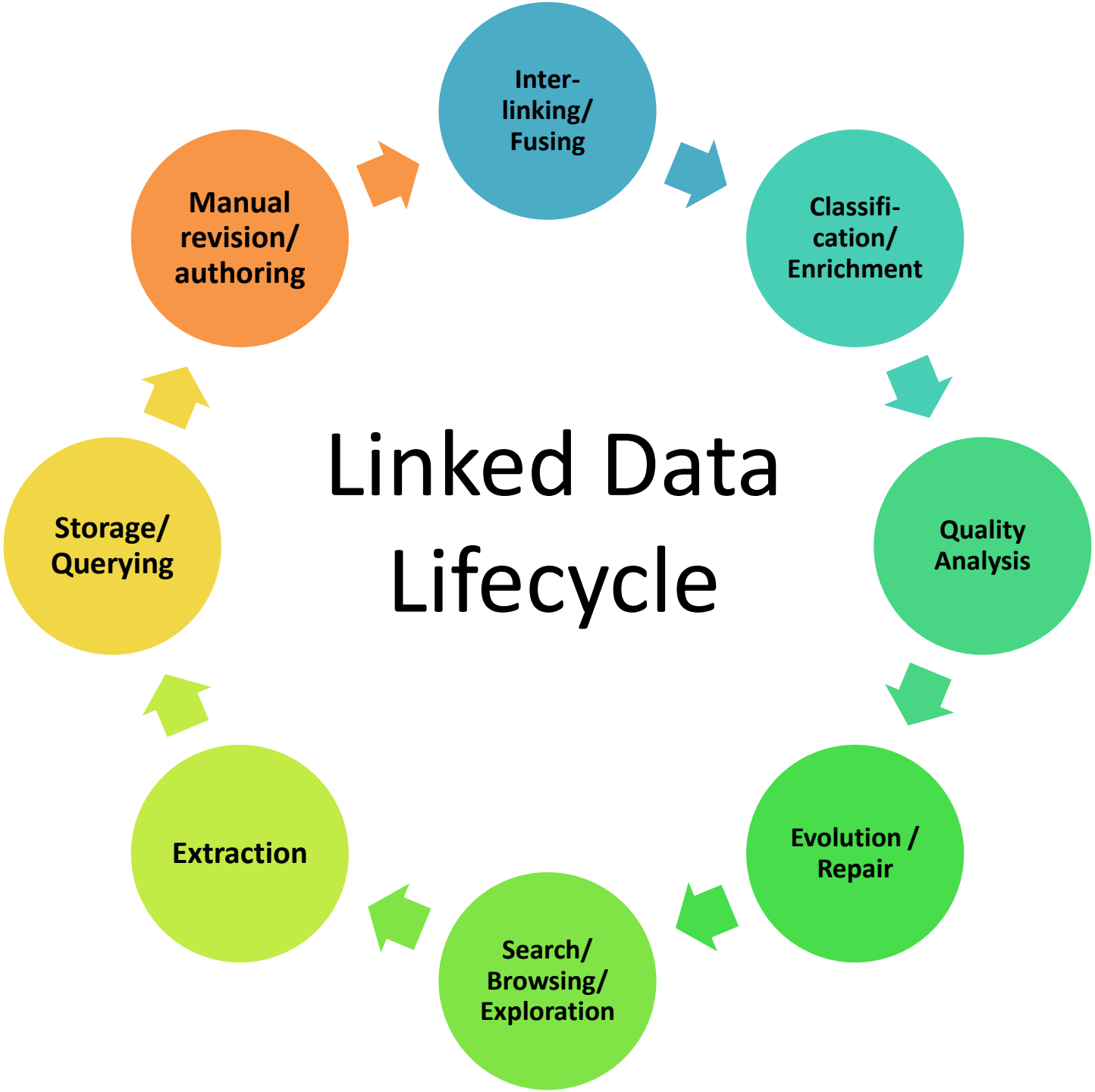
1. Extension of the Web with a **data commons** (25B facts)
2. vibrant, global RTD **community**
3. **Industrial uptake** begins (e.g. BBC, Thomson Reuters, Eli Lilly, NY Times, Facebook, Google, Yahoo)
4. Emerging **governmental adoption** in sight
5. Establishing Linked Data as **a deployment path** for the Semantic Web.

Challenges

1. **Coherence:** Relatively few, expensively maintained links
2. **Quality:** partly low quality data and inconsistencies
3. **Performance:** Still substantial penalties compared to relational
4. **Data consumption:** large-scale processing, schema mapping and data fusion still in its infancy
5. **Usability:** Missing direct end-user tools and ***network effect***

These issues are closely related **and should ultimately lead to an ecosystem of interlinked knowledge!**


Sussex Reading Lists
St. Andrews Resource Lists
NDL subjects
t4gm
RAMEAU SH
lobid Organi-sations
Audio-scrobbler (DBTune)
MySpace (DBTune)
Moseley Folk
GTAA
Plymouth Reading Lists
The Open Library
NTU Resource Lists
LIBRIS
LCSH
lobid Resources
ECS South-ampton EPrints
Ulm
DB Tropes
Magna-tune
Music Brainz (Data Incubator)
Music Brainz (zitgist)
Man-chester Reading Lists
The Open Library (Talis)
PSH
Gem. Norm-datei
UB Mann-heim
ECS South-ampton
RISKS
biz. data. gov.uk
John Peel (DB Tune)
FanHubz
Surge Radio
Discogs (Data In-cubator)
Music Brainz (DBTune)
Jamendo
Poké-pédia
Linked LCCN
VIAF
P20
RDF Book Mashup
DEPLOY
RESEX
Popula-tion (En-AKTing)
research data.gov .uk
EUTC Produc-tions
Last.fm Artists (DBTune)
Last.FM (rdfize)
classical (DB Tune)
Pokedex
PBAC
MARC Codes List
semantic web.org
Pisa
Wiki
Eurécom
NHS (EnAKTing)
Mortality (En-AKTing)
education data.gov .uk
OpenEI
BBC Program mes
BBC Music
Chronic-ling America
Event-Media
Linked MDB
NSZL Catalog
Semantic Crunch Base
Lotico
Revyu
SW Dog Food
ECS (RKB Explorer)
OAI
Budapest
Energy (En-AKTing)
CO2 (En-AKTing)
Ord-nance Survey
Openly Local
BBC Wildlife Finder
Recht-spraak. nl
Tele-graphis
New York Times
URI Burner
flickr wrappr
Open Calais
Good-win Family
RDF ohloh
BibBase
DBLP (L3S)
DBLP (RKB Explorer)
ACM
IRIT
IBM
legislation .gov.uk
UK Post-codes
statistics data.gov .uk
LOIUS
Taxon Concept
Geo Names
World Fact-book (FUB)
Freebase
iServe
VIVO UF
VIVO Indiana
VIVO Cornell
DBLP (FU Berlin)
IEEE
CiteSeer
RAE2001
New-castle
Roma
ESD stan-dards
reference data.gov .uk
Linked Data for Intervals
transport data.gov .uk
NASA (Data Incu-bator)
Fishes of Texas
Geo Species
Uberblic
DBpedia
Project Guten-berg (FUB)
OS
data dcs
STW
GESIS
Course-ware
ePrints
dotAC
CORDIS
ERA
UN/ LOCODE
The London Gazette
TWC LOGD
Eurostat
Euro-stat (FUB)
Geo Linked Data (es)
UMBEL
lingvoj
dbpedia lite
YAGO
Daily Med
TCM Gene DIT
SIDER
STITCH
OBO
Pub Chem
KEGG Drug
LAAS
KISTI
JISC
NSF
GovTrack
rdfabout US SEC
Linked Sensor Data (Kno.e.sis)
riese
Open Cyc
Lexvo
totl.net
Linked CT
Uni Path-way
Drug Bank
Medi Care
Disea-some
ChEBI
Reactome
KEGG Cpd
KEGG Pathway
KEGG Glycan
Semantic XBRL
rdfabout US Census
EUNIS
Twarql
WordNet (VUA)
Linked Open Numbers
Taxo-nomy
UniProt
Pfam
PDB
CAS
HGNC
KEGG Enzyme
KEGG Reaction
Climbing
Linked GeoData
WordNet (W3C)
Cornetto
UniRef
PRO-SITE
ProDom
SGD
Chem2 Bio2RDF
Homolo Gene
Airports
Product DB
UniSTS
UniParc
Affy-metrix
PubMed
Gene Ontology
GeneID
MGI
Gen Bank
OMIM
InterPro
Media
Geographic
Publications
User-generated content
Government
Cross-domain
Life sciences
As of September 2010


Linked Data Lifecycle
Inter-linking/ Fusing
Classifi-cation/ Enrichment
Quality Analysis
Evolution / Repair
Search/ Browsing/ Exploration
Extraction
Storage/ Querying
Manual revision/ authoring


Extraction

Extraction

From **unstructured** sources

- NLP, text mining, annotation

From **semi-structured** sources

- DBpedia, LinkedGeoData, SCOVO/DataCube

From **structured** sources

- RDB2RDF

Transforming Wikipedia into an Knowledge Base

extract structured information from Wikipedia

& make this information available on the Web as LOD:

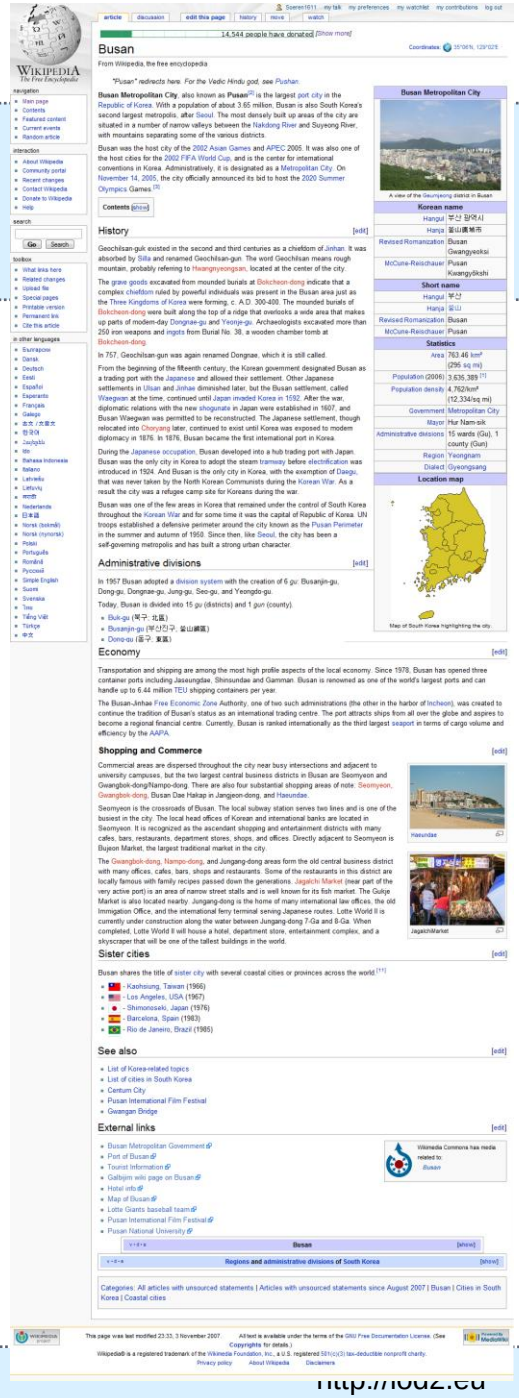
- **ask sophisticated queries against Wikipedia** (e.g. universities in brandenburg, mayors of elevated towns, soccer players),
- **link other data sets** on the Web to Wikipedia data
- Represents a **community consensus**

Recently launched DBpedia Live transforms **Wikipedia into a structred knowledge base**

Creating Knowledge
out of Interlinked Data

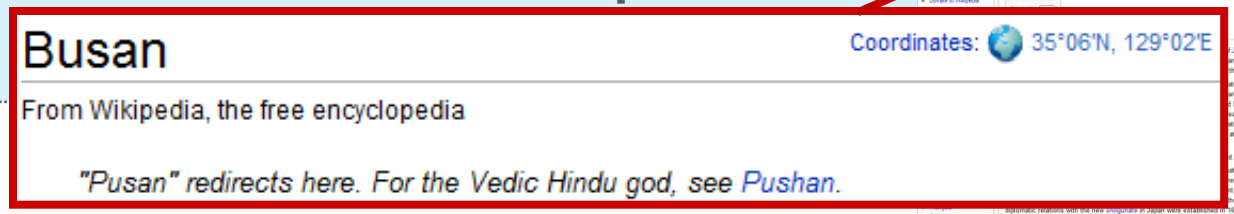
Structure in Wikipedia

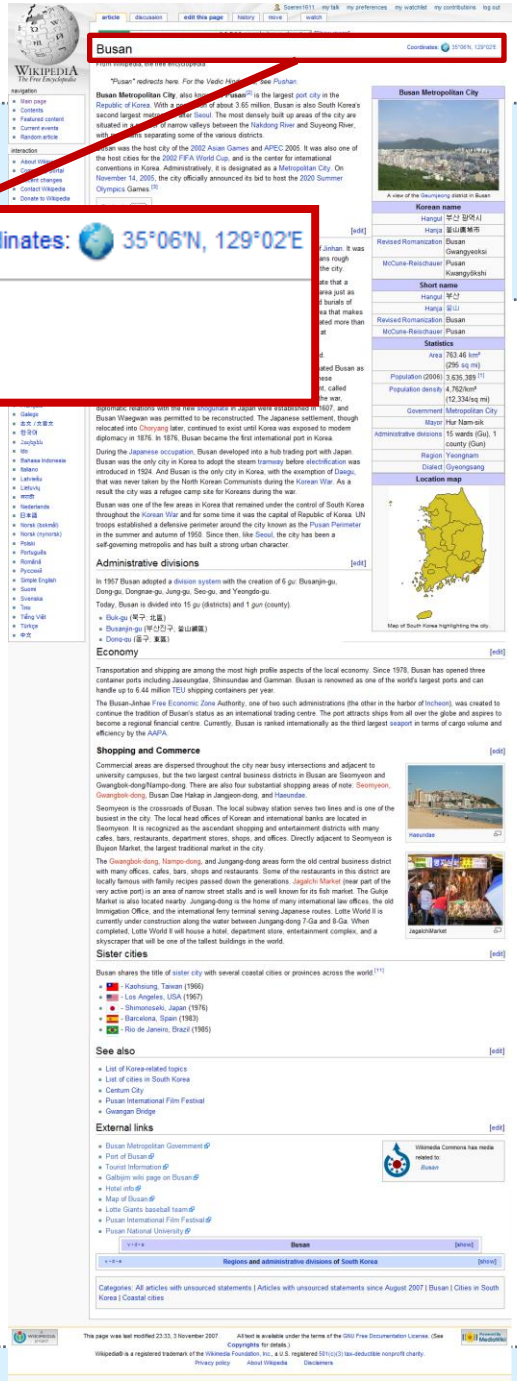
- Title
- Abstract
- Infoboxes
- Geo-coordinates
- Categories
- Images
- Links
- other language versions
- other Wikipedia pages
- To the Web
- Redirects
- Disambiguations



Creating Knowledge
out of Interlinked Data

Structure in Wikipedia

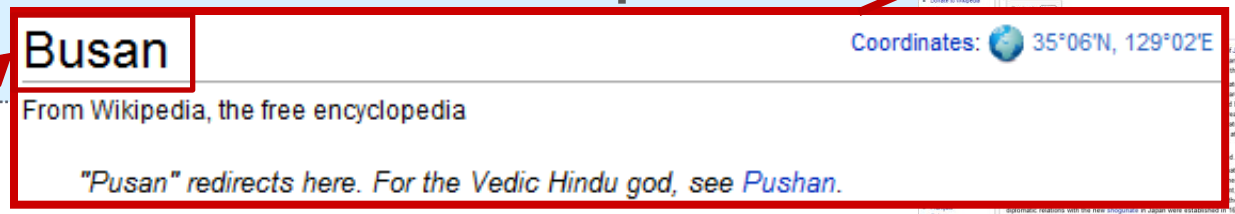




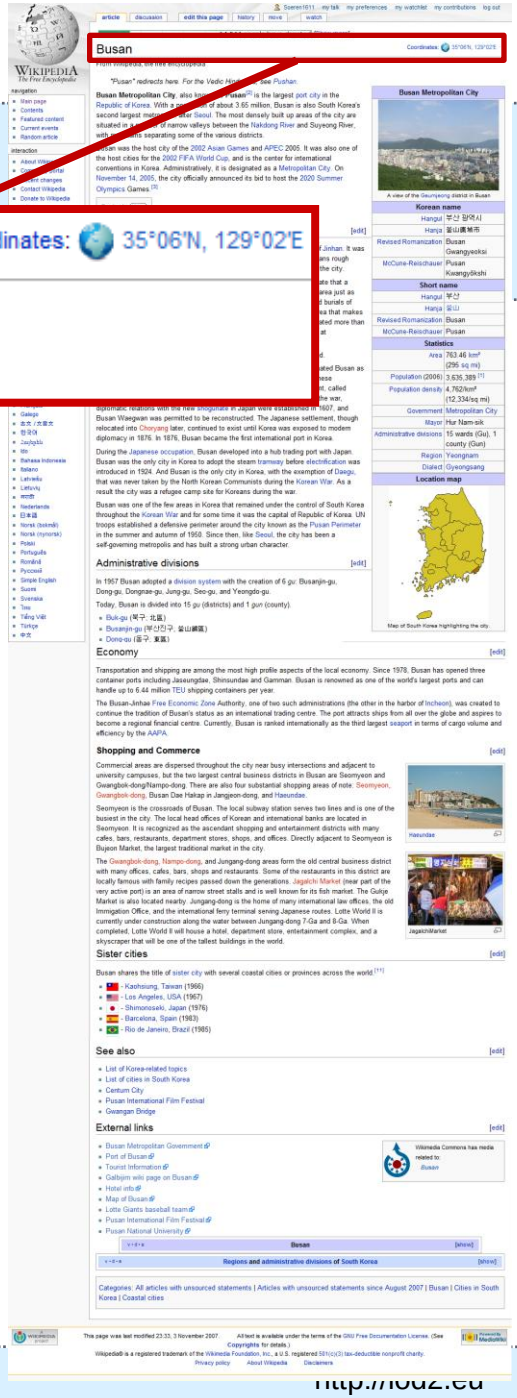
- Title
- Abstract
- Infoboxes
- Geo-coordinates
- Categories
- Images
- Links
- other language versions
- other Wikipedia pages
- To the Web
- Redirects
- Disambiguations

Creating Knowledge
out of Interlinked Data

Structure in Wikipedia



- Title
- Abstract
- Infoboxes
- Geo-coordinates
- Categories
- Images
- Links
- other language versions
- other Wikipedia pages
- To the Web
- Redirects
- Disambiguations



Creating Knowledge
out of Interlinked Data

Structure in Wikipedia

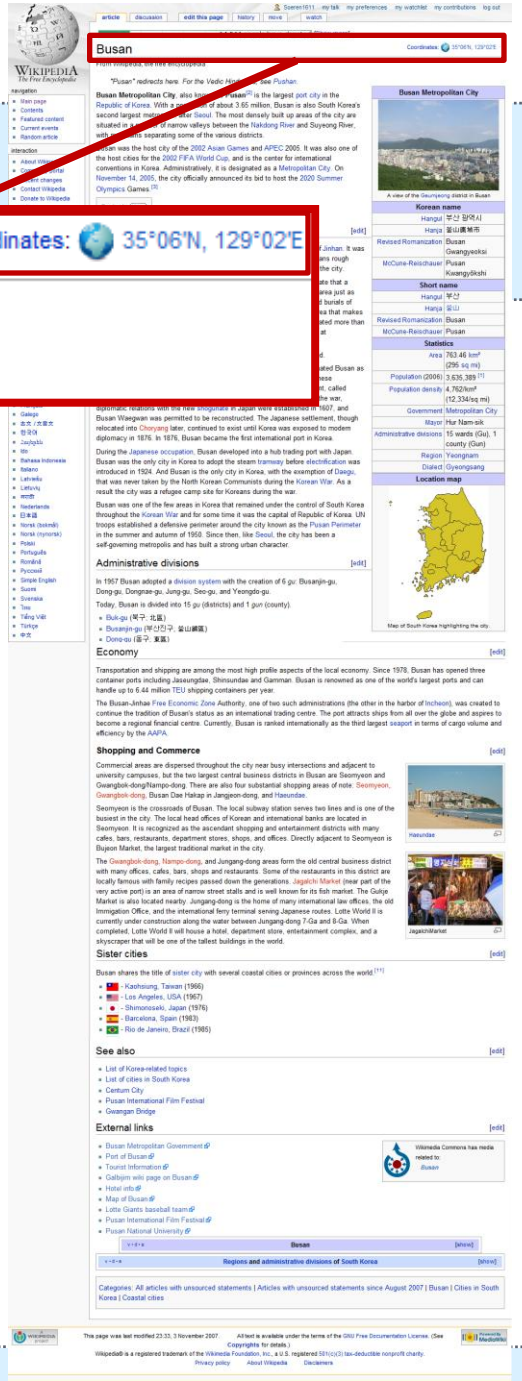
Busan

From Wikipedia, the free encyclopedia

"Pusan" redirects here. For the Vedic Hindu god, see Pushan.

Coordinates: 35°06'N, 129°02'E

- Title
- Abstract
- Infoboxes
- Geo-coordinates
- Categories
- Images
- Links
- other language versions
- other Wikipedia pages
- To the Web
- Redirects
- Disambiguations



Creating Knowledge
out of Interlinked Data

Structure in Wikipedia

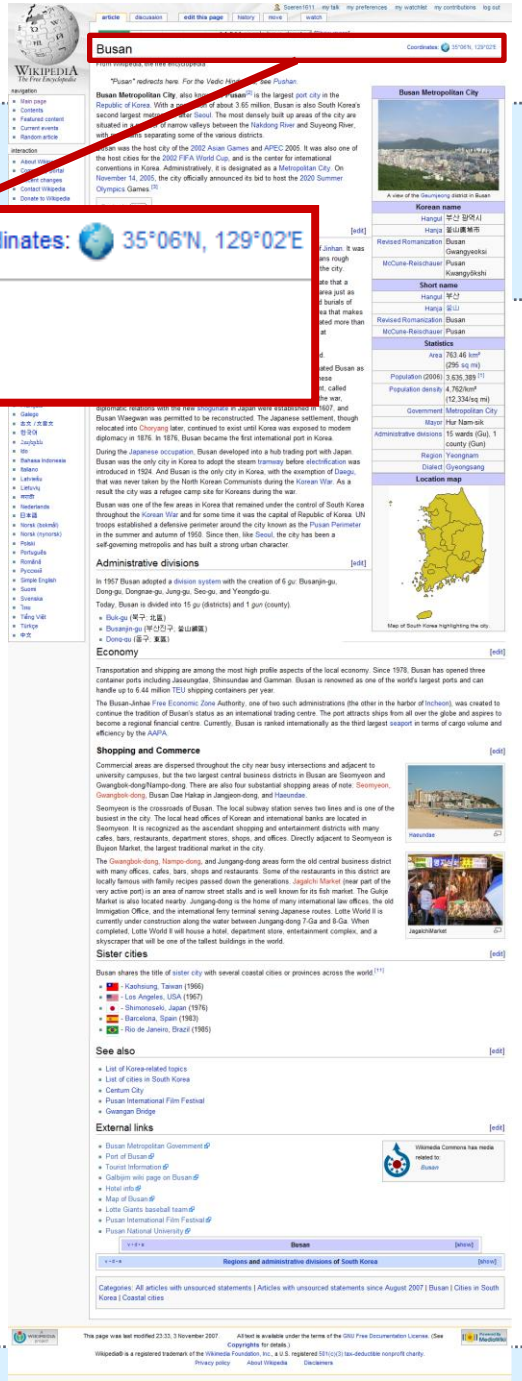
Busan

Coordinates: 35°06'N, 129°02'E

From Wikipedia, the free encyclopedia

"Pusan" redirects here. For the Vedic Hindu god, see Pushan.

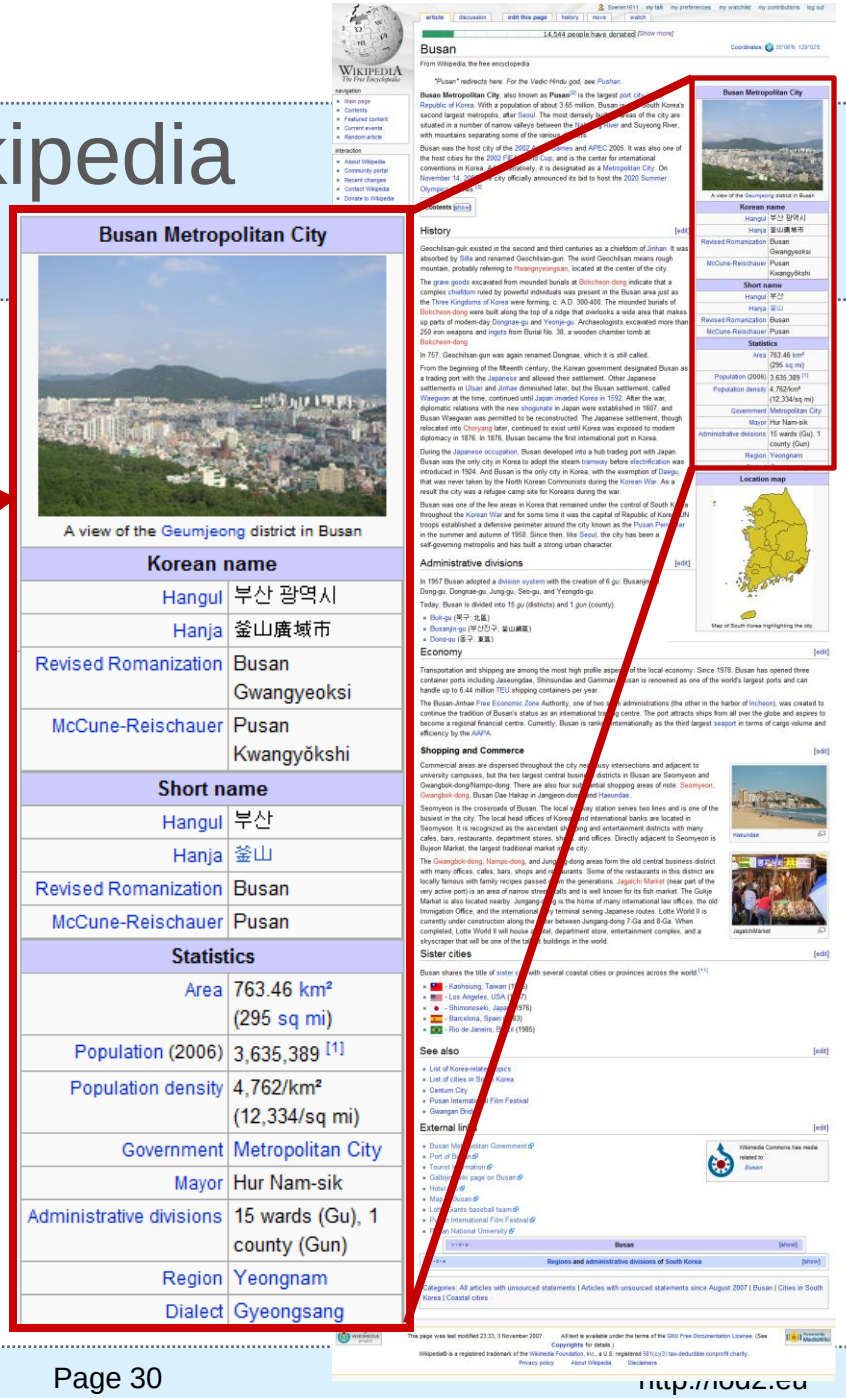
- Title
- Abstract
- Infoboxes
- Geo-coordinates
- Categories
- Images
- Links
- other language versions
- other Wikipedia pages
- To the Web
- Redirects
- Disambiguations



Creating Knowledge
out of Interlinked Data

Structure in Wikipedia

- Title
- Abstract
- Infoboxes
- Geo-coordinates
- Categories
- Images
- Links
- other language versions
- other Wikipedia pages
- To the Web
- Redirects
- Disambiguations



Busan Metropolitan City

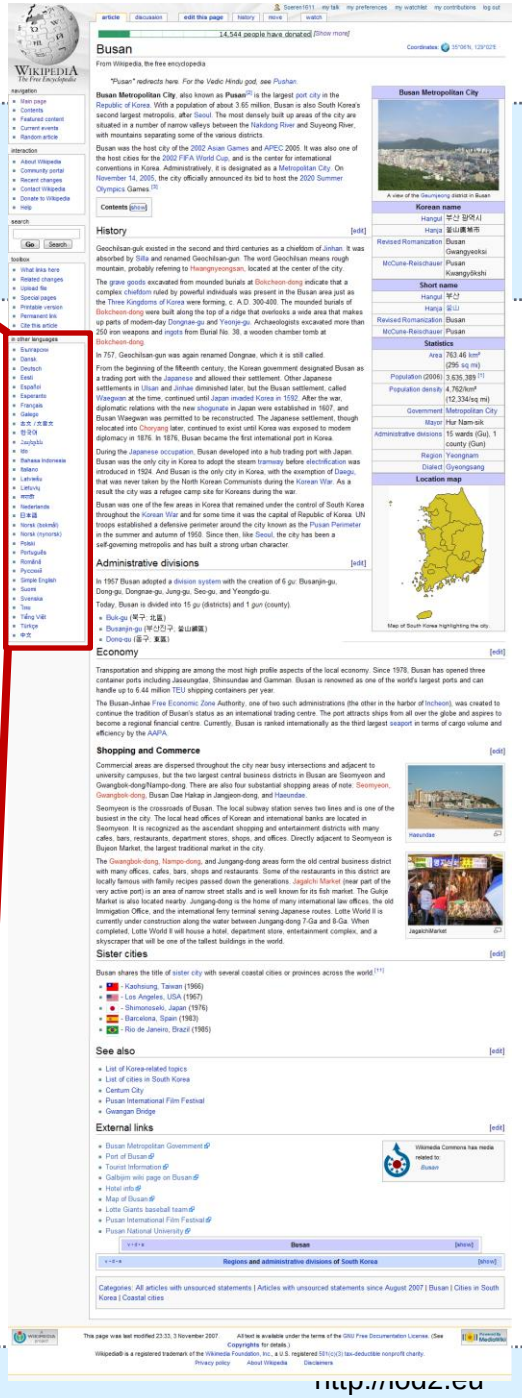
A view of the Geumjeong district in Busan

Korean name	
Hangul	부산 광역시
Hanja	釜山廣域市
Revised Romanization	Busan Gwangyeoksi
McCune-Reischauer	Pusan Kwangyŏkshi
Short name	
Hangul	부산
Hanja	釜山
Revised Romanization	Busan
McCune-Reischauer	Pusan
Statistics	
Area	763.46 km² (295 sq mi)
Population (2006)	3,635,389 [1]
Population density	4,762/km² (12,334/sq mi)
Government	Metropolitan City
Mayor	Hur Nam-sik
Administrative divisions	15 wards (Gu), 1 county (Gun)
Region	Yeongnam
Dialect	Gyeongsang

Creating Knowledge
out of Interlinked Data

Structure in Wikipedia

- Title
- Abstract
- Infoboxes
- Geo-coordinates
- Categories
- Images
- Links
- other language versions
- other Wikipedia pages
- To the Web
- Redirects
- Disambiguations



Creating Knowledge
out of Interlinked Data

Structure in Wikipedia

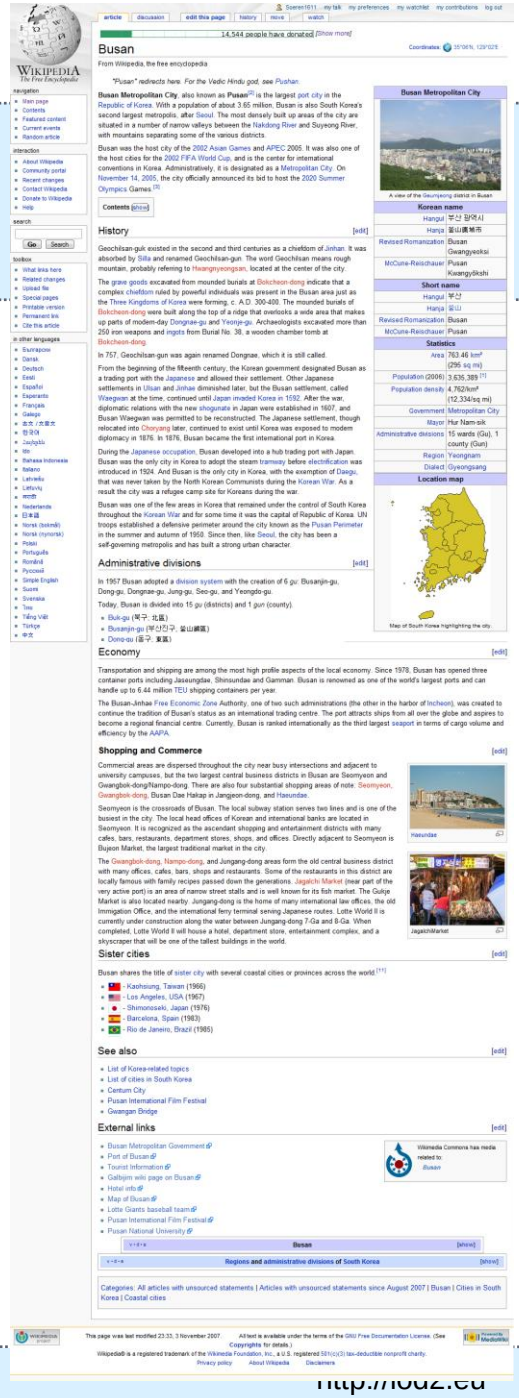
- Title
- Abstract
- Infoboxes
- Geo-coordinates
- Categories
- Images
- Links
- other language versions
- other Wikipedia pages
- To the Web
- Redirects
- Disambiguations

Categories: Busan | Cities in South Korea | Coastal cities

Creating Knowledge
out of Interlinked Data

Structure in Wikipedia

- Title
- Abstract
- Infoboxes
- Geo-coordinates
- Categories
- Images
- Links
- other language versions
- other Wikipedia pages
- To the Web
- Redirects
- Disambiguations



LOD2
Creating Knowledge
out of Interlinked Data

DBpedia

Infobox templates

Wikitext-Syntax

```
{{Infobox Korean settlement
| title          = Busan Metropolitan City
| img            = Busan.jpg
| imgcaption     = A view of the [[Geumjeong]] district in Bu
| hangul         = 부산 광역시
...
| area_km2          = 763.46
| pop            = 3635389
| popyear        = 2006
| mayor          = Hur Nam-sik
| divs           = 15 wards (Gu), 1 county (Gun)
| region         = [[Yeongnam]]
| dialect        = [[Gyeongsang]]
}}
```

http://dbpedia.org/resource/Busan

RDF representation

```
dbp:Busan   dbpp:title      "Busan Metropolitan City"
dbp:Busan   dbpp:hangul     "부산 광역시"@Hang
dbp:Busan   dbpp:area_km2   "763.46"^^xsd:float
dbp:Busan   dbpp:pop        "3635389"^^xsd:int
dbp:Busan   dbpp:region     dbp:Yeongnam
dbp:Busan   dbpp:dialect    dbp:Gyeongsang
...
```

Busan Metropolitan City

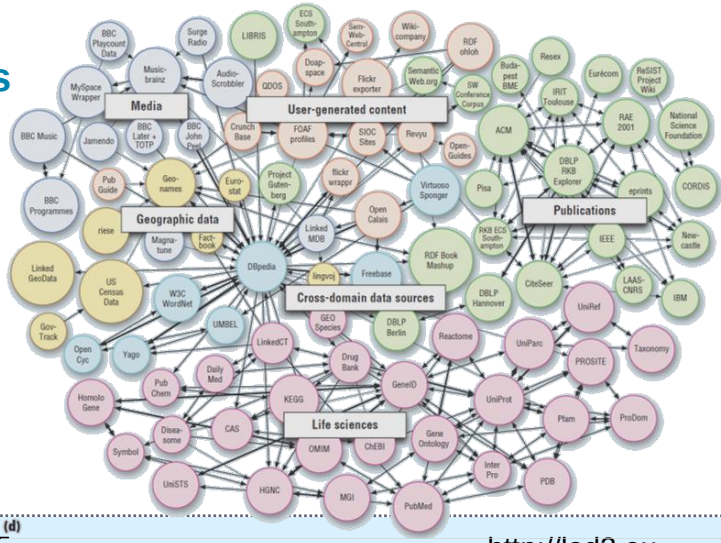
A view of the Geumjeong district in Busan

Korean name	
Hangul	부산 광역시
Hanja	釜山廣域市
Revised Romanization	Busan Gwangyeoksi
McCune-Reischauer	Pusan Kwangyŏkshi
Short name	
Hangul	부산
Hanja	釜山
Revised Romanization	Busan
McCune-Reischauer	Pusan
Statistics	
Area	763.46 km² (295 sq mi)
Population (2006)	3,635,389 [1]
Population density	4,762/km² (12,334/sq mi)
Government	Metropolitan City
Mayor	Hur Nam-sik
Administrative divisions	15 wards (Gu), 1 county (Gun)
Region	Yeongnam
Dialect	Gyeongsang

LOD2
Creating Knowledge
out of Interlinked Data

A vast multi-lingual, multi-domain knowledge base

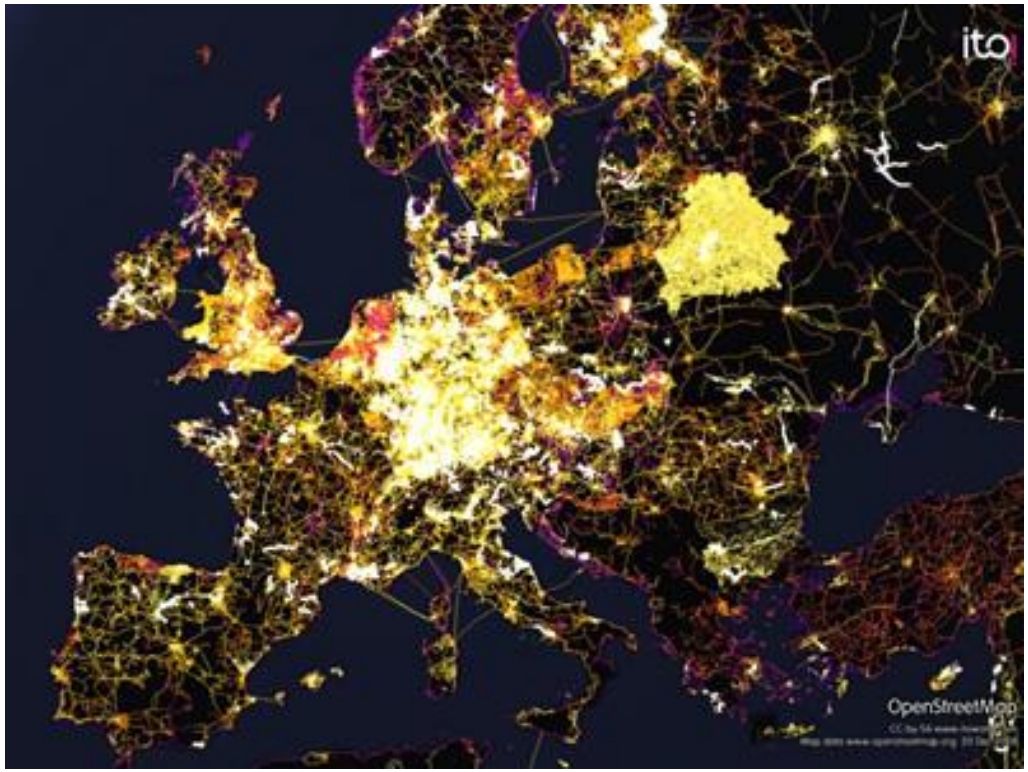
- **descriptions of ca. 3.4 million things** (1.5 million classified in a consistent ontology), including 312K *persons*, 413K *places*, 94K *music albums*, 49K *films*, 15K *video games*, 140K *organizations*, 146K *species*, 5K *diseases*
- labels & abstracts in up to 92 different languages
- altogether **>1 billion pieces of information** (i.e. RDF triples): 257M from English edition, 766M from other language editions
  - **substantial impact** in science, technology and society
  - became a **central interlinking hub** on the Data Web
  - Scientific publications attracted more than **500 citations**
  - More than 15.000 monthly visits on DBpedia.org, numerous press articles, blog posts …
  - Ecosystem of commercial and community applications: ThomsonReuters, BBC, Neofonie, Openlink, Faviki …



OpenStreetMap<->LinkedGeoData

OpenStreetMap is an incredible source of spatial data
LinkedGeoData makes this data available as Linked Data

- Mappings to DBpedia, Geonames
- SPARQL Endpoint
- REST API – Linked Data Services



Users	309465
GPS points	1993482813
Nodes	795615514
Ways	66102033
Relations	751638

This **faceted Linked Geo Data browser** is based on data obtained from the OpenStreetMap project (released under CC-BY-SA) and was developed by AKSW research group.

LinkedGeoData.org

Facets | **Instances** | **Search:** powered by Namefinder

Class

[+] amenity(397)
atm (5) | bank (8) | bench (101) | bicycle_parking (17) | bicycle_rental (1) | biergarten (2) | cafe (11) | car_wash (1) | cinema (2) | circus (1) | clock (1) | dentist (1) | doctors (2) | fast_food (16) | fountain (9) | grit_bin (11) | hospital (1) | kindergarten (5) | library (1) | nightclub (2) | parking (15) | pharmacy (8) | post_box (18) | post_office (1) | pub (7) | recycling (24) | restaurant (34) | school (7) | shelter (1) | taxi (5) | telephone (43) | theatre (1) | toilets (11) | university (2) | vending_machine (18) | veterinary (1) | waste_basket (2) | water (1) |

[+] highway(109)
[+] historic(6)
[+] leisure(10)
[+] man_made(60)
[+] natural(69)
[+] railway(76)
[+] shop(60)
[+] sport(4)
[+] tourism(21)
[+] waterway(1)

Properties

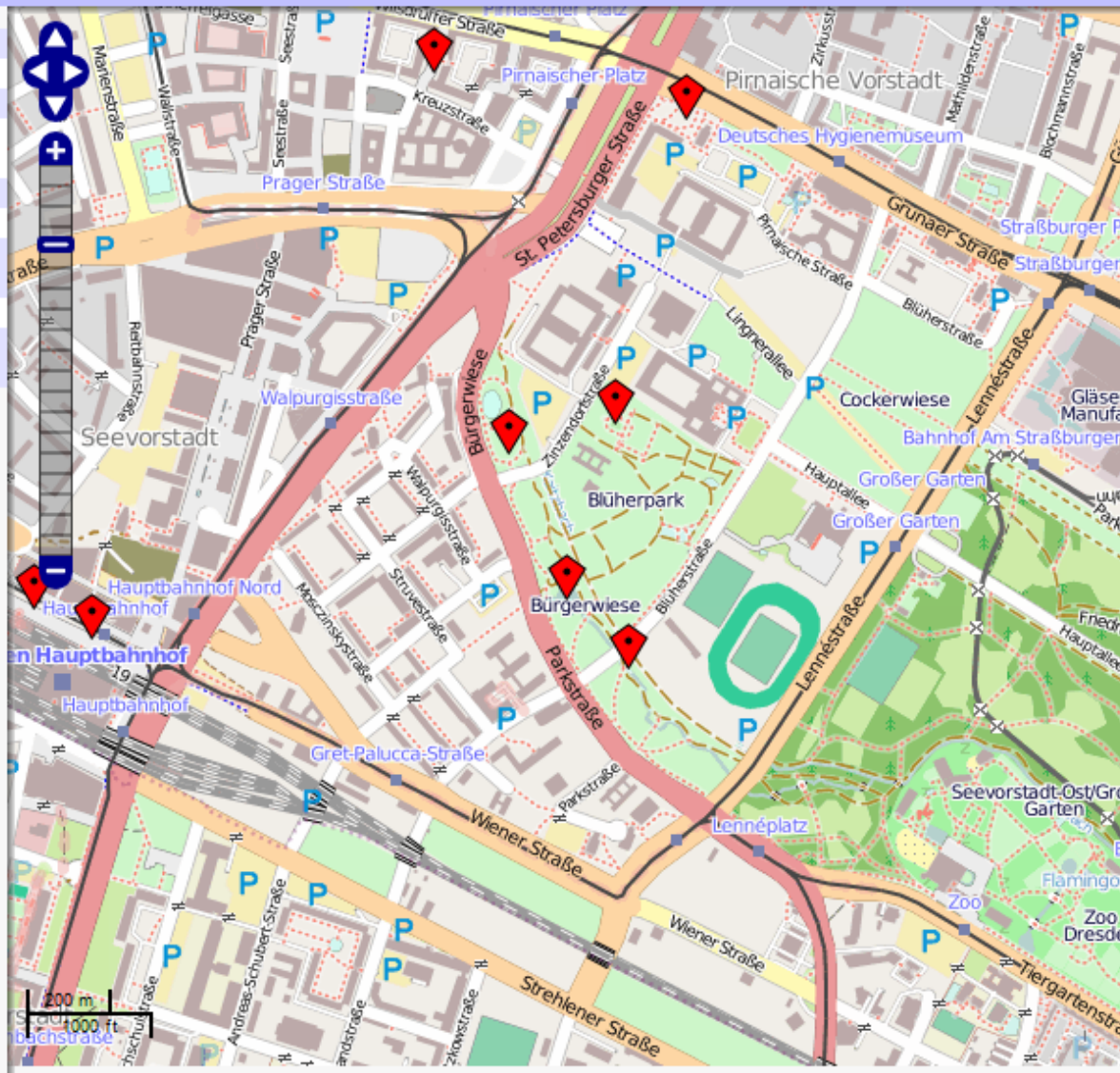
[+] cuisine(18)
[+] place(1)

deselect

Query took 141ms.

1. amenity: fountain
2. amenity: fountain
3. amenity: fountain
4. **Badende**
   amenity: fountain
5. amenity: fountain
6. amenity: fountain
7. **Mozart-Brunnen**
   amenity: fountain
8. amenity: fountain
9. **Gänsediebbrunnen**
   amenity: fountain

Query took 136ms.



Extraction Relational Data

Many different approaches (D2R, Virtuoso RDF Views, Triplify, …)

No agreement on a formal semantics of RDF2RDF mapping

- LOD readiness, SPARQL-SQL translation

W3C RDB2RDF WG

Tool	Triplify	D2RQ	Virtuoso RDF Views
Technology	Scripting languages (PHP)	Java	Whole middleware solution
SPARQL endpoint	-	X	X
Mapping language	SQL	RDF based	RDF based
Mapping generation	Manual	Semi-automatic	Manual
Scalability	Medium-high (but no SPARQL)	Medium	High

Extraction Challenges

From **unstructured** sources

- Deploy existing NLP approaches (OpenCalais, Ontos API)
- Develop standardized, LOD enabled interfaces between NLP tools (NLP2RDF)

From **semi-structured** sources

- Efficient bi-directional synchronization

From **structured** sources

- Declarative syntax and semantics of data model transformations (W3C WG RDB2RDF)

Orthogonal challenges

- Using LOD as background knowledge
- Provenance

Storage and Querying

RDF Data Management

Still by a factor 5-50 slower than relational data management

Performance increases steadily

Comprehensive, well-supported open-soure and commercial implementations are available:

- OpenLink's Virtuoso (os+commercial)
- Big OWLIM (commercial), Swift OWLIM (os)
- Talis (hosted)
- Bigdata (distributed)
- Allegrograph (commercial)
- Mulgara (os)

Adaptive SPARQL Query Cache

RDF Graph:

```
…
JensLehmann      rdf:type        foaf:person
JensLehmann      foaf:knows      SoerenAuer
JensLehmann      foaf:birthDay   “10.4.1981”
SebastianTramp   rdf:type        foaf:person
SebastianTramp   foaf:knows      SoerenAuer
…
```

Adaptive SPARQL Query Cache

RDF Graph:

```
…
JensLehmann      rdf:type        foaf:person
JensLehmann      foaf:knows      SoerenAuer
JensLehmann      foaf:birthDay   “10.4.1981”
SebastianTramp   rdf:type        foaf:person
SebastianTramp   foaf:knows      SoerenAuer
…
```

Graph pattern (SPARQL core):

```
?person   rdf:type       foaf:person
?person   foaf:knows     http://aksw.org/SoerenAuer
```

Adaptive SPARQL Query Cache

RDF Graph:

```
…
JensLehmann      rdf:type        foaf:person
JensLehmann      foaf:knows      SoerenAuer
JensLehmann      foaf:birthDay   “10.4.1981”
SebastianTramp   rdf:type        foaf:person
SebastianTramp   foaf:knows      SoerenAuer
…
```

Graph pattern (SPARQL core):

```
?person   rdf:type       foaf:person
?person   foaf:knows     http://aksw.org/SoerenAuer
```

Adaptive SPARQL Query Cache

RDF Graph:

```
…
JensLehmann      rdf:type        foaf:person
JensLehmann      foaf:knows      SoerenAuer
JensLehmann      foaf:birthDay   “10.4.1981”
SebastianTramp   rdf:type        foaf:person
SebastianTramp   foaf:knows      SoerenAuer
…
```

Graph pattern (SPARQL core):

```
?person   rdf:type      foaf:person
?person   foaf:knows    http://aksw.org/SoerenAuer
```

Question: Which query results are affected by which updates?

Adaptive SPARQL Query Cache

RDF Graph:

```
…
JensLehmann      rdf:type        foaf:person
JensLehmann      foaf:knows      SoerenAuer
JensLehmann      foaf:birthDay   “10.4.1981”
SebastianTramp   rdf:type        foaf:person
SebastianTramp   foaf:knows      SoerenAuer
…
```

Graph pattern (SPARQL core):

```
?person   rdf:type      foaf:person
?person   foaf:knows    http://aksw.org/SoerenAuer
```

Question: Which query results are affected by which updates?

Theorem: *Query results only change, when updated triples match one of the triple patterns.* Formal proof wrt. SPARQL algebra in [1].

[1] Michael Martin, Jörg Unbehauen, Sören Auer: Improving the Performance of Semantic Web Applications with SPARQL Query Result Caching. In proceedings of 7th Extended Semantic Web Conference (ESWC2010).

Adaptive SPARQL Query Cache

RDF Graph:

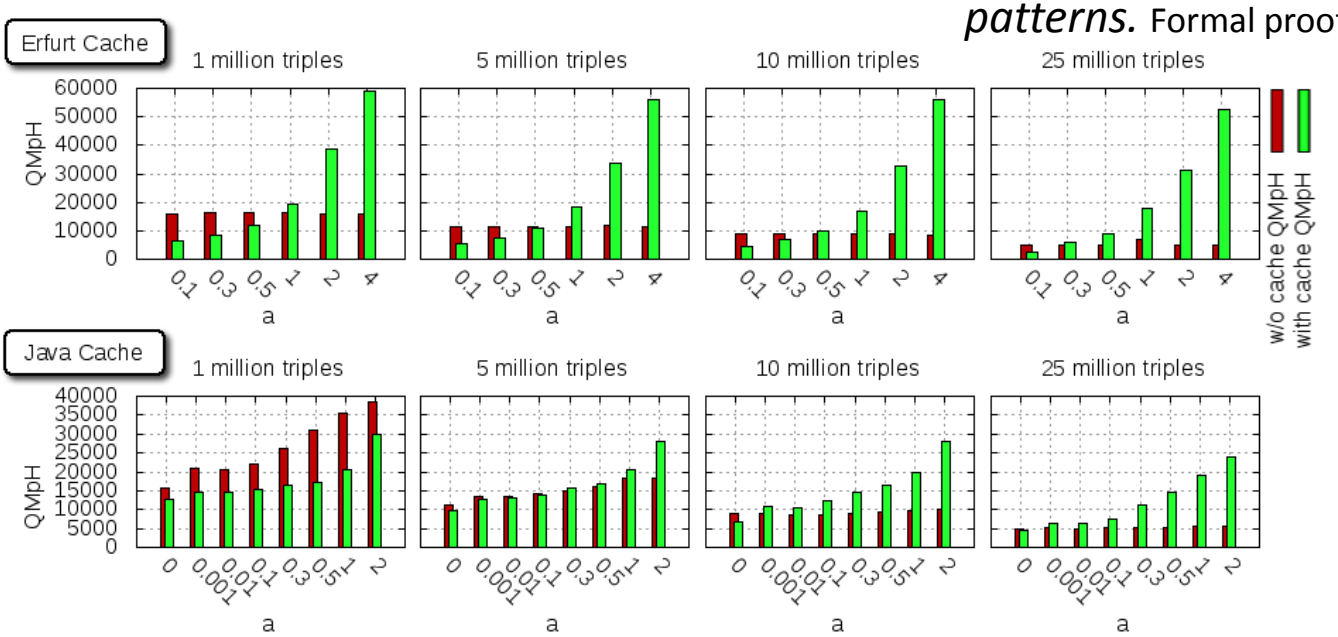
```
...
JensLehmann      rdf:type       foaf:person
JensLehmann      foaf:knows     SoerenAuer
JensLehmann      foaf:birthDay  "10.4.1981"
SebastianTramp   rdf:type       foaf:person
SebastianTramp   foaf:knows     SoerenAuer
...
```

Graph pattern (SPARQL core):

```
?person  rdf:type     foaf:person
?person  foaf:knows   http://aksw.org/SoerenAuer
```

Question: Which query results are affected by which updates?

Theorem: *Query results only change, when updated triples match one of the triple patterns.* Formal proof wrt. SPARQL algebra in [1].

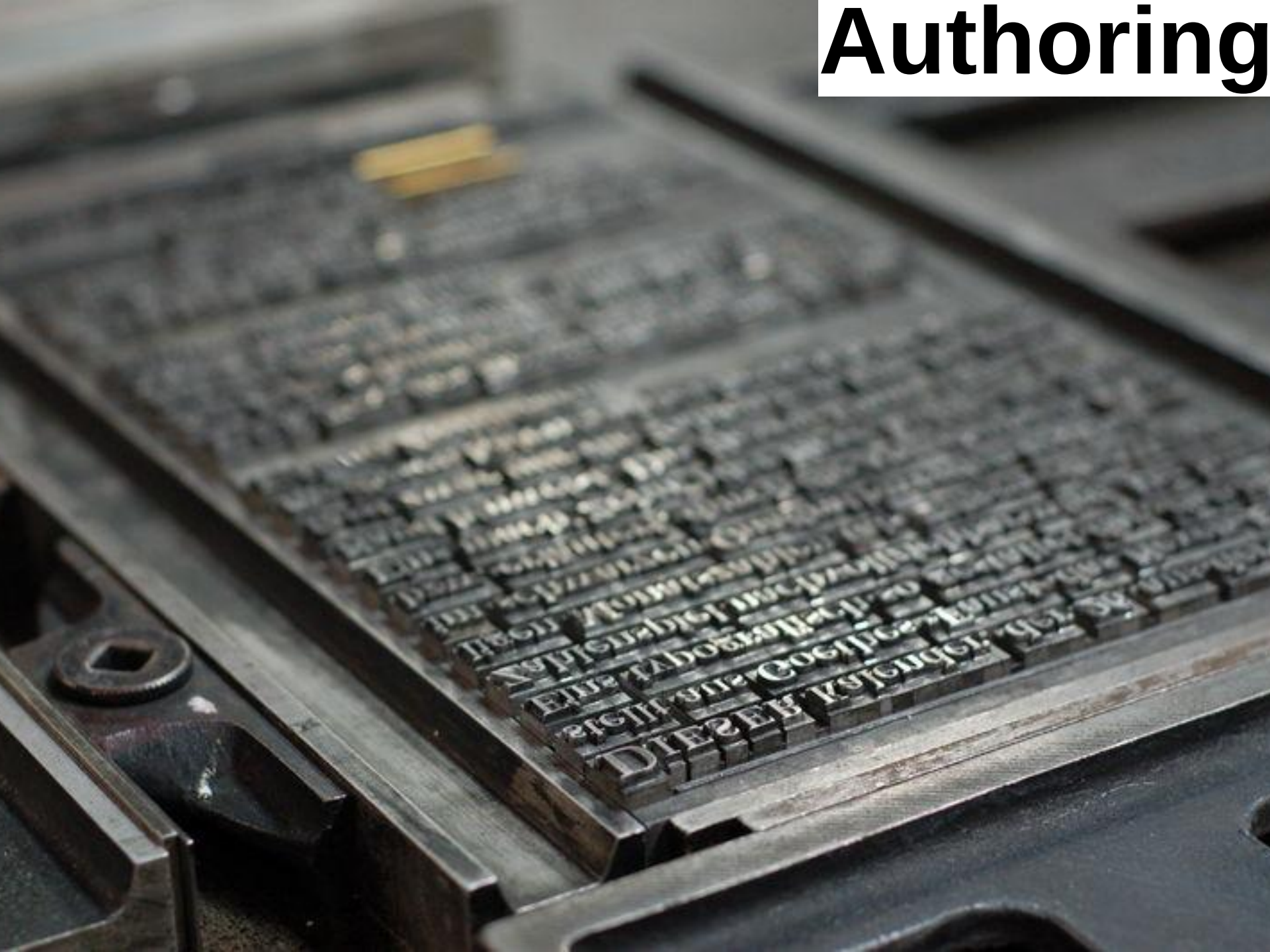


Work with **OpenLink** Ltd. (UK) on integrating the approach into the **Virtuoso** triple store and with **CWI** (NL) into **MonetDB**

[1] Michael Martin, Jörg Unbehauen, Sören Auer: Improving the Performance of Semantic Web Applications with SPARQL Query Result Caching. In proceedings of 7th Extended Semantic Web Conference (ESWC2010).

Storage and Querying Challenges

- Reduce the performance gap between relational and RDF data management
- SPARQL Query extensions
- Spatial/semantic/temporal data management
- More advanced query result caching
- View maintenance / adaptive reorganization based on common access patterns
- More realistic benchmarks



Authoring

Two Kinds of Semantic Wikis

1. Semantic (Text) Wikis

- Authoring of semantically annotated texts

2. Semantic Data Wikis

- Direct authoring of structured information (i.e. RDF, RDF-Schema, OWL)

Two Kinds of Semantic Wikis

1. Semantic (Text) Wikis

- Authoring of semantically annotated texts

article | discussion | edit | history | protect | delete | move | watch | ref

Editing Australia

```
This is a sample page that is used in the SMW documentation
process.  This page represents an "individual" - in this case, the
country of Australia.  ([[:Special:ExportRDF/{{PAGENAME}}|RDF
Rendering]])

According to
https://www.cia.gov/cia/publications/factbook/print/as.html,
Australia [[hasPopulation:=20,264,082]had an estimated population
of 20,264,082]] as of July, 2006.

The country of Australia is [[locatedOn::Australian_Continent]]

[[Category:Country]]
```

2. Semantic Data Wikis

- Direct authoring of structured information (i.e. RDF, RDF-Schema, OWL)

LOD2
Creating Knowledge
out of Interlinked Data

Two Kinds of Semantic Wikis

1. Semantic (Text) Wikis

- Authoring of semantically annotated texts

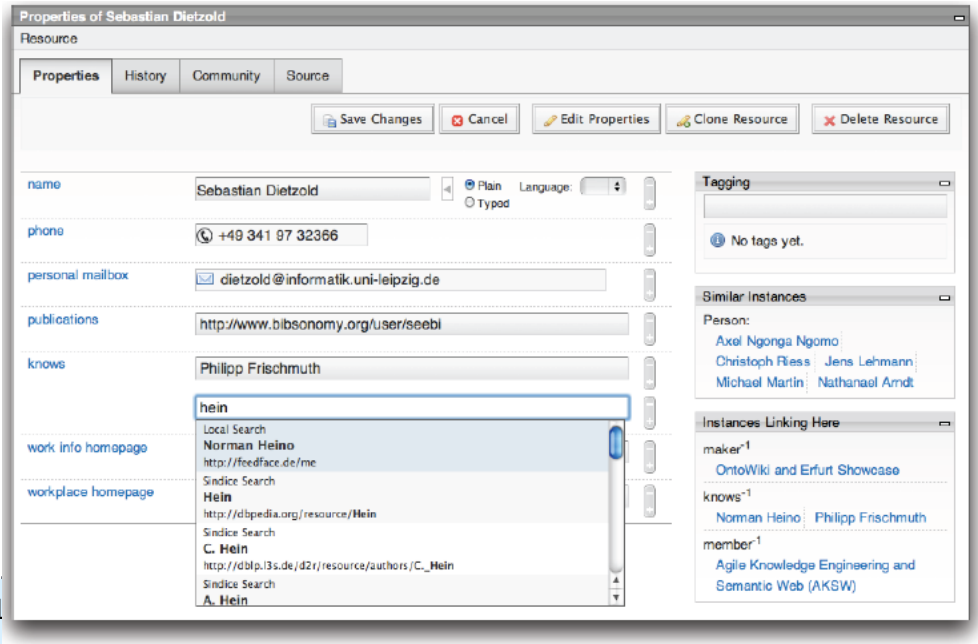
article | discussion | edit | history | protect | delete | move | watch | ref

Editing Australia

```
This is a sample page that is used in the SMW documentation
process.  This page represents an "individual" - in this case, the
country of Australia.  ([[:Special:ExportRDF/{{PAGENAME}}|RDF
Rendering]])

According to
https://www.cia.gov/cia/publications/factbook/print/as.html,
Australia [[hasPopulation:=20,264,082]] had an estimated population
of 20,264,082]] as of July, 2006.

The country of Australia is [[locatedOn::Australian_Continent]]

[[Category:Country]]
```

2. Semantic Data Wikis

- Direct authoring of structured information (i.e. RDF, RDF-Schema, OWL)



OntoWiki – a semantic data wiki

Versatile domain-independent tool

Serves as Linked Data / SPARQL endpoint on the Data Web

Open-source project hosted at Google code

Not just a Wiki UI, but a whole framework for the development of Semantic Web applications

Developed in PHP based on the Zend framework

Very active developer and user community

More than 500 downloads monthly

Large number of use cases

Management of multimedia content for a record label in the UK

OntoWiki (Sebastian Tramp)

User Extras Help Debug

Search for Resources

POWERED BY VIRTUOSO

Knowledge Bases

Edit View

OntoWiki System Configuration

Caucasus Spiders

Navigation: Checklist

Edit View Type

Search in Navigation

Caucasus Spiders

- Ukraine
- Russia
- Turkey
- Iran
- Moldavia

Resource List

View

Instances | History | Community | Source

		published in	location
1.	33167: Harpactea golovatchi Original Record	Dunin (1992)	Shvanidzor
2.	33169: Harpactea eskovi Original Record	Dunin (1992)	Shnokh
3.	33168: Harpactea eskovi Original Record	Dunin (1992)	Ozun
4.	33172: Harpactea eskovi Original Record	Dunin (1992)	Savarlo
5.	33173: Harpactea nenilini Original Record	Dunin (1992)	Kuris
6.	33170: Harpactea eskovi Original Record	Dunin (1992)	Vanadzor
7.	33175: Harpactea deelemanae Original Record	Dunin (1992)	Kajaran
8.	33174: Harpactea nenilini Original Record	Dunin (1992)	Shishkert
9.	40155: Macaroeris nidicolens Original Record	Logunov & Rakov (1998)	Mukhtadir
10.	33178: Harpactea azerbajdzhanica Original Record	Dunin (1992)	Aurora

rdf:type	Original Record
label	33168: Harpactea eskovi
comment	type series 92
location	Ozun
species name	Harpactea eskovi
published in	Dunin (1992)

« First ‹ Previous 1 2 3 Next › Last » Search returned 5179 results. Show me: 10 50 100 ... all Query execution took 68 ms.

Explore Tags

View

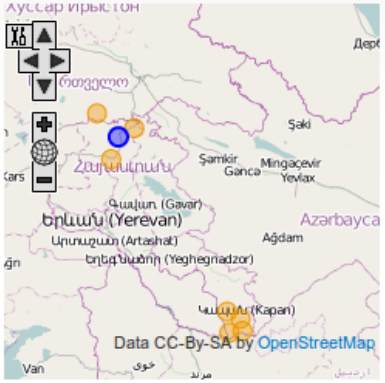
within

Temirgoe · Rossk. · Evdokimovsk · Kergez Mt. · Bazardyuzyu Mts.

Minimap

Edit View Zoom

Search for location



Show Properties

rdf:type · label · comment · location · species name · published in · within

Filter

Search in list

Add complex Filter Clear

Active Filters:

- rdf:type equals Original Record

Properties of Philipp Frischmuth

Resource

http://philipp.frischmuth24.de/id/me View Resource

Properties | History | Community | Source

Data was found for the given URI. 28 statements were added.

ns1:to	http://philipp.frischmuth24.de/pingback/ping/
type	Person
seeAlso	http://feeds.delicious.com/rss/pfrischmuth
sameAs	5562 id490bd762-0352-40e8-8559-3726c31bbd8a
birthday	1984-10-09
current project	http://pingback.aksw.org/
depiction	
firstName	Philipp
gender	male
holds account	pfrischmuth pfrischmuth
homepage	http://www.frischmuth24.de
knows	me Sebastian Tramp
personal mailbox	mailto:frischmuth@informatik.uni-leipzig.de mailto:philipp@frischmuth24.de
name	Philipp Frischmuth
nickname	Phil

Tagging

AKSW OntoWiki

Tag removed

Similar Instances

Person

Norman Heino • me • Sebastian Tramp • AxelCNgongaNgomo • ChristophRiess • ...

Instances Linking Here

knows^{-1}

Norman Heino • Sebastian Tramp

maker^{-1}

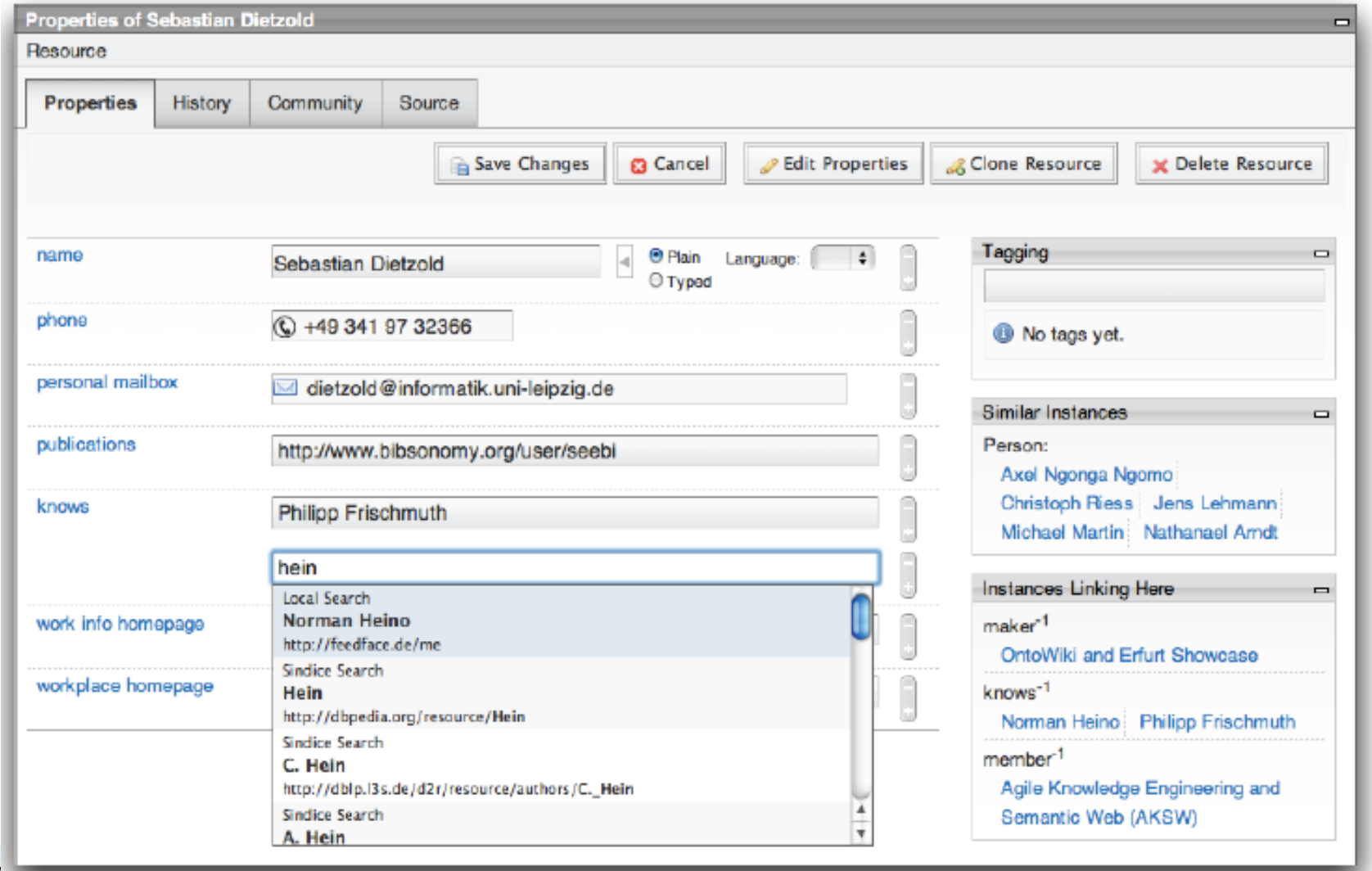
OntoWiki and Erfurt Showcase • Erfurt API • OntoWiki • Semantic Pingback • Triplify

member^{-1}

Agile Knowledge Engineering and Semantic Web (AKSW)

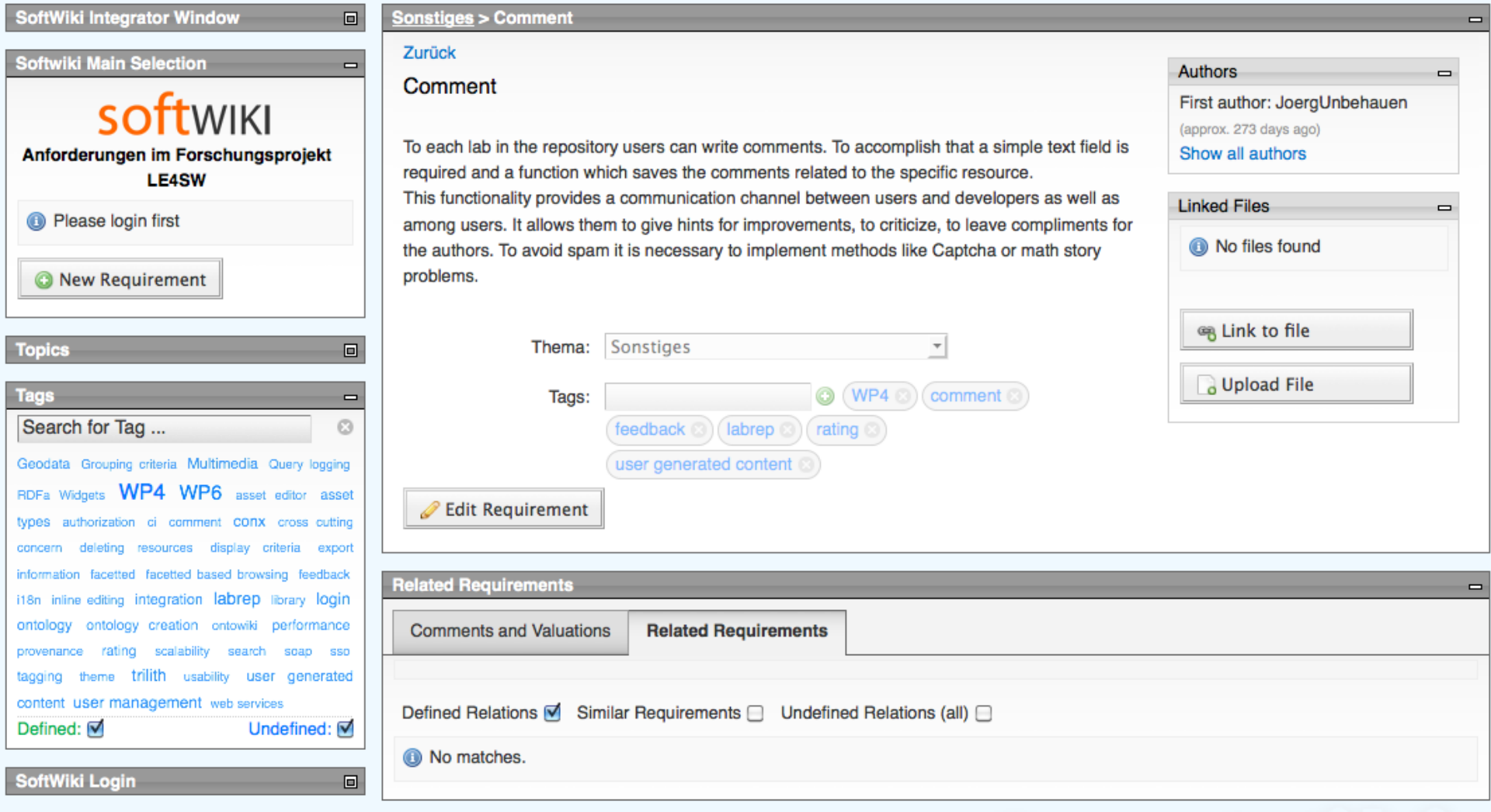
LOD2
Creating Knowledge
out of Interlinked Data

RDFauthor in OntoWiki



LOD2
Creating Knowledge
out of Interlinked Data

OntoWiki: Supporting Requirements engineering



Linking

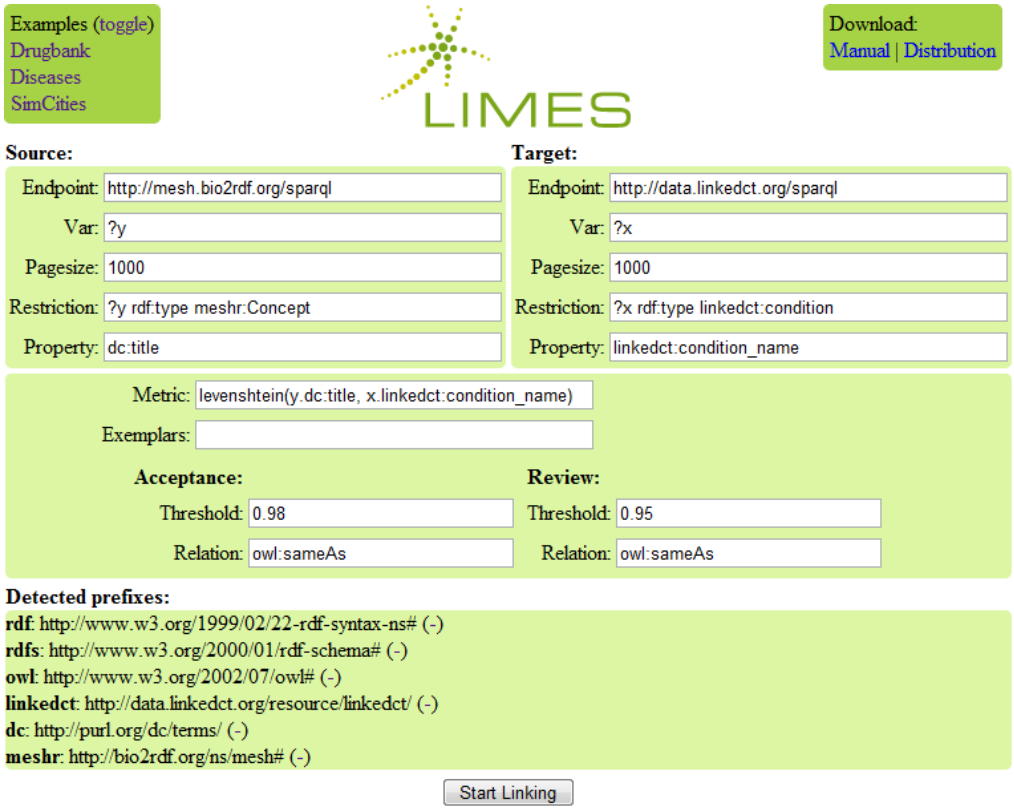
LOD Linking

Manual:

- SILK
- LIMES

Semi-Automatic:

- SAIM (soon)

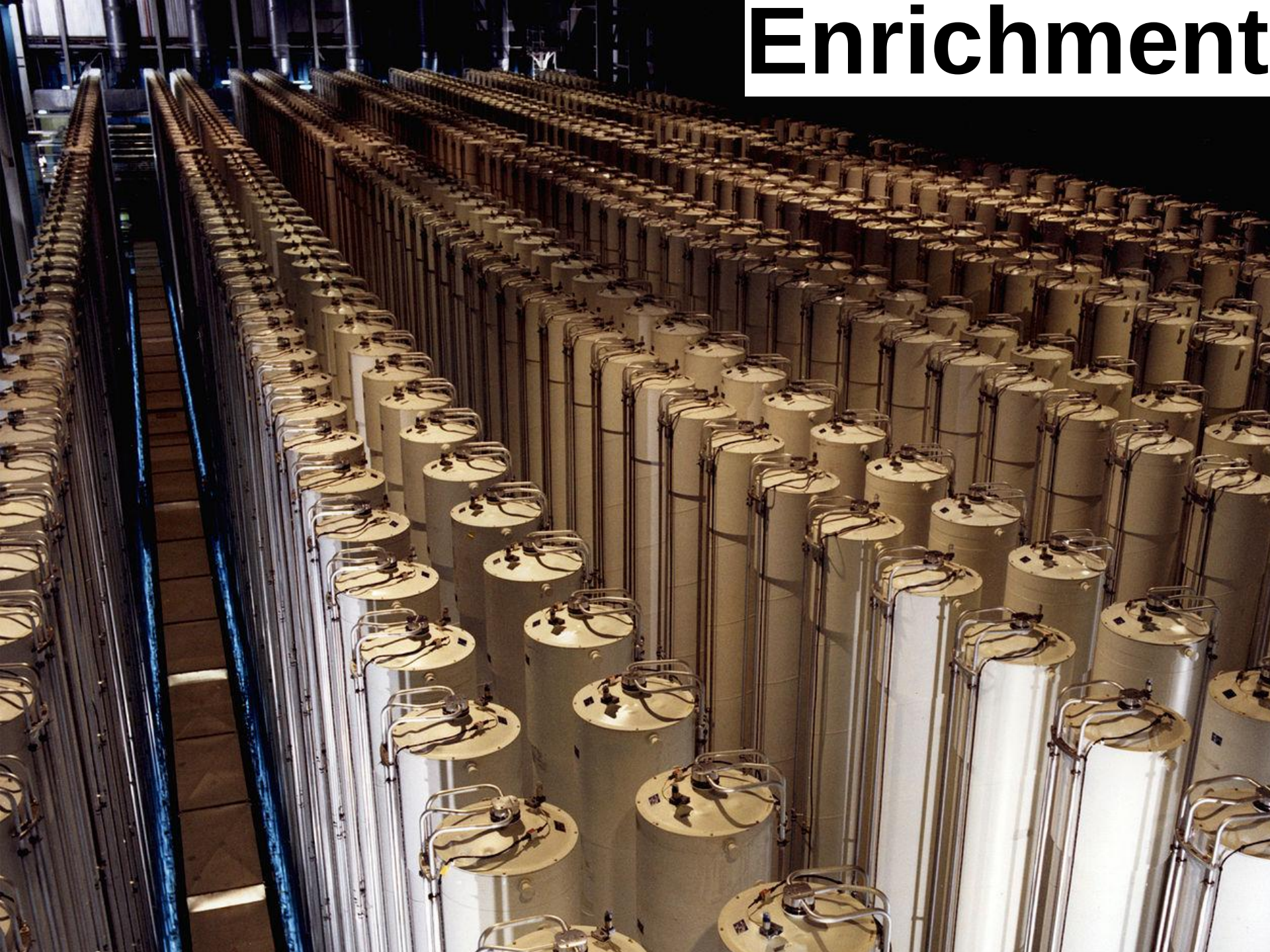


Interlinking Challenges

Only 5% of the information on the Data Web is actually linked

- Make sense of work in the de-duplication/record linkage literature
- Consider the open world nature of Linked Data
- Use LOD background knowledge
- Zero-configuration linking
- Explore active learning approaches, which integrate users in a feedback loop
- Maintain a 24/7 linking service: Linked Open Data Around-The-Clock project (LATC-project.eu)

Enrichment



Enrichment & Repair

ORE
ONTOLOGY REPAIR
AND ENRICHMENT

ORE (**O**ntology **R**epair and **E**nrichment) tool allows to improve an OWL ontology by fixing inconsistencies & making suggestions for adding further axioms.

- **Ontology Debugging:** OWL reasoning to detect inconsistencies and satisfiable classes + detect the most likely sources for the problems. user can create a repair plan, while maintaining full control.
- **Ontology Enrichment:** uses the DL-Learner framework to suggest definitions & super classes for existing classes in the KB. works if instance data is available for harmonising schema and data.

http://aksw.org/Projects/ORE

Improving Linked Data **Quality** by Ontology Learning

Given:

- **Background knowledge** base
- Positive and negative **examples** (example = individual in ontology)

Goal:

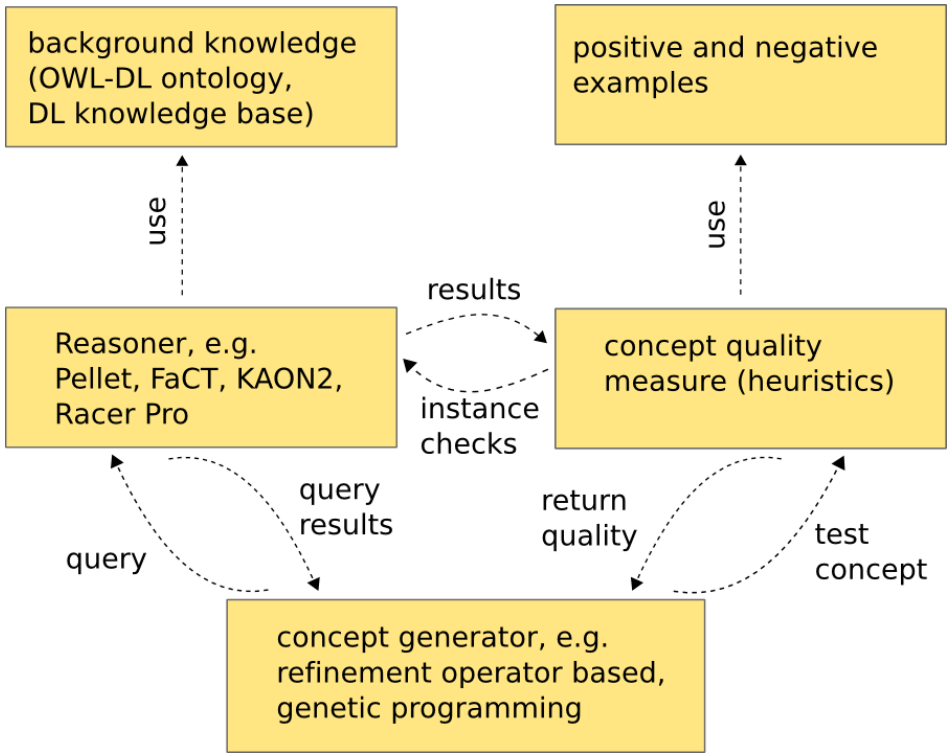
Find an OWL **Class Expression** / DL concept which

- **covers** as **many positive examples** as possible
- **covers** as **few negative examples** as possible

Concept C covers example a <=> a is instance of C

Analogous problem can be defined for logic programs => Inductive Logic Programming

Supervised Machine Learning Task



Framework for Supervised Machine Learning for OWL & Description Logics

Application areas:

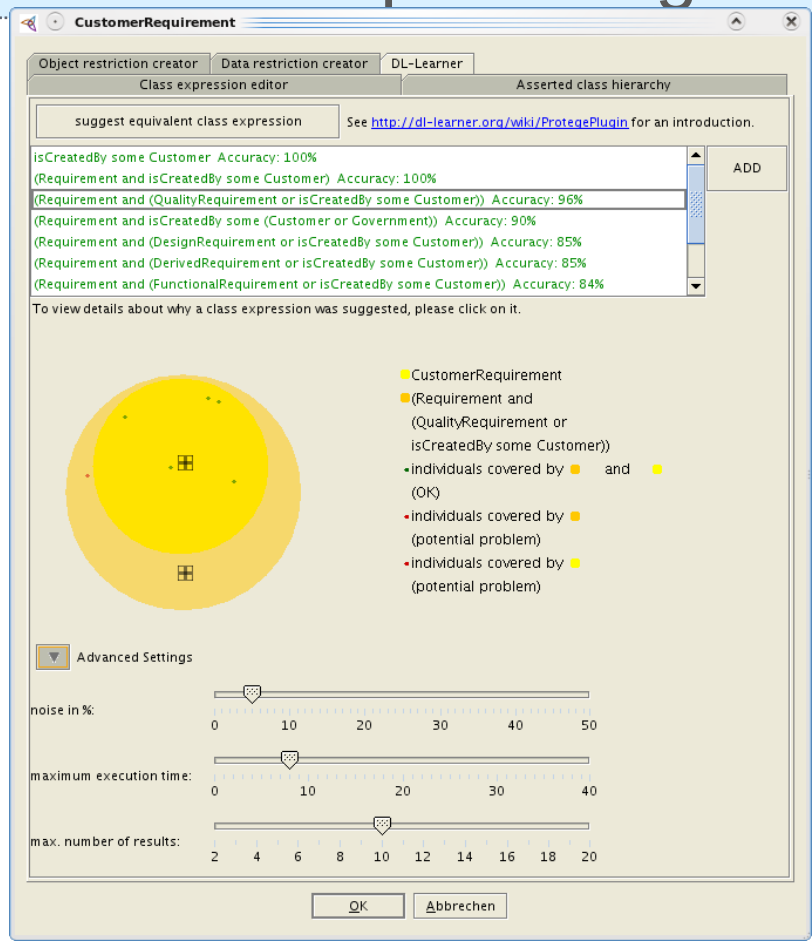
- Ontology Engineering: learn axioms for existing classes, PlugIns for Protégé and OntoWiki
- recommendation/navigation, e.g. music based on last songs heard or navigation suggestions in large knowledge bases (Dbpedia navigation)
- “Classical” Machine Learning, e.g. predicting Carcinogenesis

Works on OWL Files and SPARQL endpoints

Supports different reasoner interfaces

Accessible via command-line, GUI, web service

Implemented in Java, available as open-source,



Hellmann, Lehmann, Auer: Learning of OWL Class Descriptions on Very Large Knowledge Bases. Int. Journal on Semantic Web & Information Systems (IJSWIS), Vol. 5, Issue 2, April-July 2009, ISSN: 1552-6283.
Jens Lehmann, Sören Auer: Class Expression Learning for Ontology Engineering. To appear in Journal of Web Semantics (JWS).

Quality Analysis

Linked Data Quality Analysis

Quality on the Data Web is varying a lot

- Hand crafted or expensively curated knowledge base (e.g. DBLP, UMLS) vs. extracted from text or Web 2.0 sources (DBpedia)

Research Challenge

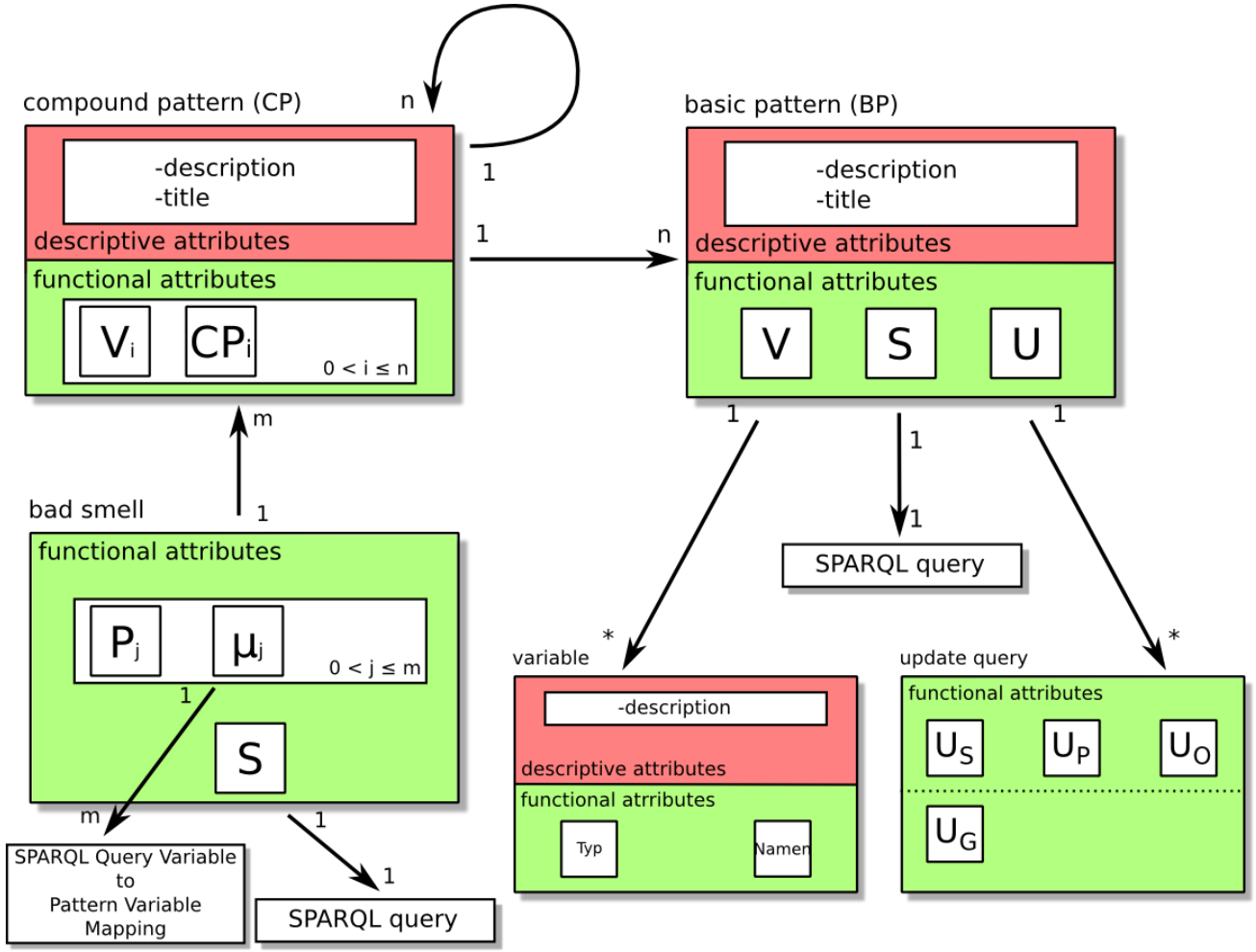
- Establish measures for assessing the authority, provenance, reliability of Data Web resources

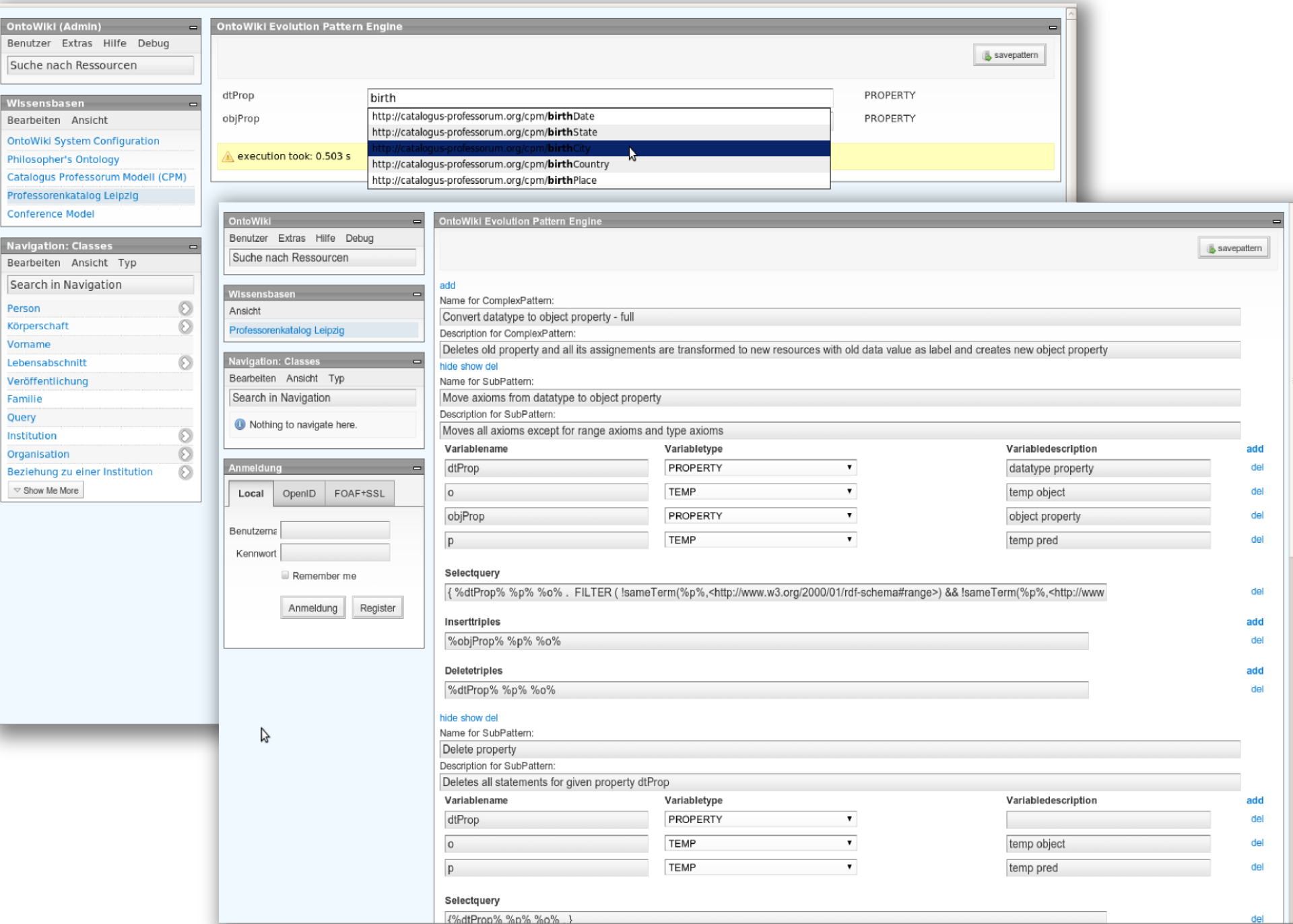
Evolution

EvoPat – Pattern based KB Evolution

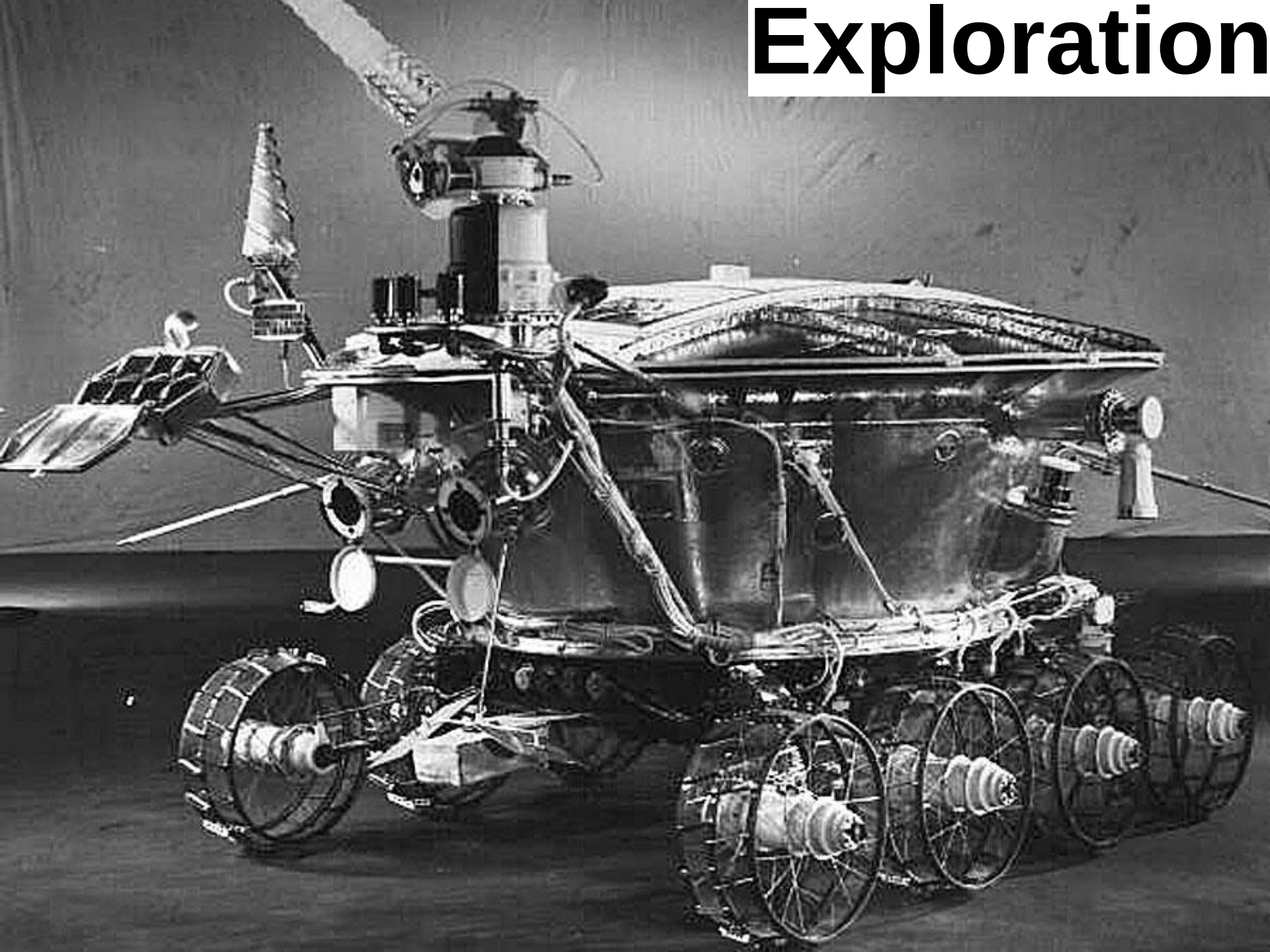
- unified method, for both data evolution and ontology refactoring.
- modularized, declarative definition of evolution patterns is relatively simple compared to an imperative description of evolution
  - allows domain experts and knowledge engineers to amend the ontology structure and modify data with just a few clicks
- Combined with RDF representation of evolution patterns and their exposure on the Linked Data Web, EvoPat facilitates the development of an evolution pattern ecosystem
  - patterns can be shared and reused on the Data Web.
- declarative definition of **bad smells and corresponding evolution patterns** promotes the (semi-)automatic improvement of information quality.

Evolution Patterns




OntoWiki (Admin)
Benutzer Extras Hilfe Debug
Suche nach Ressourcen
Wissensbasen
Bearbeiten Ansicht
OntoWiki System Configuration
Philosopher's Ontology
Catalogus Professorum Modell (CPM)
Professorenkatalog Leipzig
Conference Model
Navigation: Classes
Bearbeiten Ansicht Typ
Search in Navigation
Person
Körperschaft
Vorname
Lebensabschnitt
Veröffentlichung
Familie
Query
Institution
Organisation
Beziehung zu einer Institution
Show Me More
OntoWiki Evolution Pattern Engine
savepattern
dtProp
birth
PROPERTY
objProp
PROPERTY
http://catalogus-professorum.org/cpm/birthDate
http://catalogus-professorum.org/cpm/birthState
http://catalogus-professorum.org/cpm/birthCity
http://catalogus-professorum.org/cpm/birthCountry
http://catalogus-professorum.org/cpm/birthPlace
execution took: 0.503 s
OntoWiki
Benutzer Extras Hilfe Debug
Suche nach Ressourcen
Wissensbasen
Ansicht
Professorenkatalog Leipzig
Navigation: Classes
Bearbeiten Ansicht Typ
Search in Navigation
Nothing to navigate here.
Anmeldung
Local
OpenID
FOAF+SSL
Benutzerna
Kennwort
Remember me
Anmeldung
Register
OntoWiki Evolution Pattern Engine
savepattern
add
Name for ComplexPattern:
Convert datatype to object property - full
Description for ComplexPattern:
Deletes old property and all its assignements are transformed to new resources with old data value as label and creates new object property
hide show del
Name for SubPattern:
Move axioms from datatype to object property
Description for SubPattern:
Moves all axioms except for range axioms and type axioms
Variablename
Variabletype
Variabledescription
add
dtProp
PROPERTY
datatype property
del
o
TEMP
temp object
del
objProp
PROPERTY
object property
del
p
TEMP
temp pred
del
Selectquery
{ %dtProp% %p% %o% . FILTER (!sameTerm(%p%,<http://www.w3.org/2000/01/rdf-schema#range>) && !sameTerm(%p%,<http://www
del
Inserttriples
add
%objProp% %p% %o%
del
Deletetriples
add
%dtProp% %p% %o%
del
hide show del
Name for SubPattern:
Delete property
Description for SubPattern:
Deletes all statements for given property dtProp
Variablename
Variabletype
Variabledescription
add
dtProp
PROPERTY
del
o
TEMP
temp object
del
p
TEMP
temp pred
del
Selectquery
{%dtProp% %p% %o% . }
del

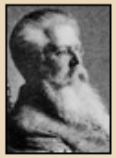
Exploration



LOD2
Creating Knowledge

Catalogus Professorum Lipsiensis




Professorenkatalog der Universität Leipzig | catalogus professorum lipsiensis

Startseite

Epochen

Fakultäten

Professoren des Tages

Rektoren und Dekane

gesamtes Verzeichnis

Volltextsuche

Hintergrund

Information in English

Prof. Dr. phil. *Levin* Ludwig Schücking

Bildquelle: Privatbesitz

Lebensdaten

geb. 29.05.1878 in Burgsteinfurt
gest. 12.10.1964 in Farchant

Lebenslauf

Studium

- 1897-1901 Studium: Neuere Sprachen und Kunstgeschichte in Freiburg, Göttingen, Berlin und München

Qualifikation

- 1904 Habilitation für Englische Philologie an der Universität Göttingen
  Titel der Arbeit: Grundzüge der Satzverknüpfung im Beowulf.
- 1901 Promotion zum Dr. phil. in Englischer Philologie an der Universität Göttingen
  Titel der Arbeit: Stoffliche Beziehungen der englischen Komödie zur italienischen bis Lilly.

Professorenkatalog Leipzig

User View Extras Help

Suchtext

Alle Wissensbasen durchsuchen

Suchen

Wissensbasen

Modell des Professorenkatalogs

Professorenkatalog der Universität Leipzig

Sprachen (ausgezeichnete Labels)

de en keine

Klassen

- Vorname (693)
- Person (1534)
  - Professor (1312)
- Veröffentlichung (1262)
- Lebensabschnitt (10299)
  - Beziehung zu einer Institution (8840)
    - Karriere (4367)
    - Amt (803)
    - Qualifikationsschrift (2385)
    - Studium (1285)
  - gesellschaftliche Funktion (1412)
- Verwandschaftsbeziehungen (1)
  - Geschwister (1)
    - Bruder (1)
- Körperschaft (423)

Person > Professor > Levin Ludwig Schücking

Resource Edit View

Eigenschaften | Karte | Kalender | Versionen | Bearbeiten | Kommentare und Bewertungen

Properties for Levin Ludwig Schücking

Nachname	Schücking
voller akademischer Titel	Prof. Dr. phil.
eine Person hat Vornamen	• Levin • Ludwig
Person ID	144
Geburtstag	1878-05-29
Geburtsstadt	Burgsteinfurt
Geburtsland	Preußen
Todestag	1964-10-12
Todesstadt	Farchant
Bild	
Bildherkunft	Privatbesitz
Links (e.g.Wikipedia)	http://de.wikipedia.org/wiki/Levin_Ludwig_Sch%C3%BCcking
Abitur Stadt	Münster
Abitur Land	Preußen
Professor hat	• 1897-1901 Studium: Neuere Sprachen und Kunstgeschichte in Freiburg,

Autoren

Erstellt vor ca. 530 Tagen durch Admin

Letzte Änderung vor ca. 253 Tagen durch ChristianAugustin

Weitere Autoren:

- FriederikeBulka
- ReginaSteiner
- ChristianRau

Ähnliche Instanzen

Professor

Georg Hermann Schnedermann | Carl Friedrich Heinrich Schaffstein | Wolfgang Otto Bernhard | Schadewaldt | Fritz Adolf Albert | Prinzhorn | Ludwig Weickmann

Show all 1312 in Main Tab »

Hierher verweisende Instanzen

ist Betreuer

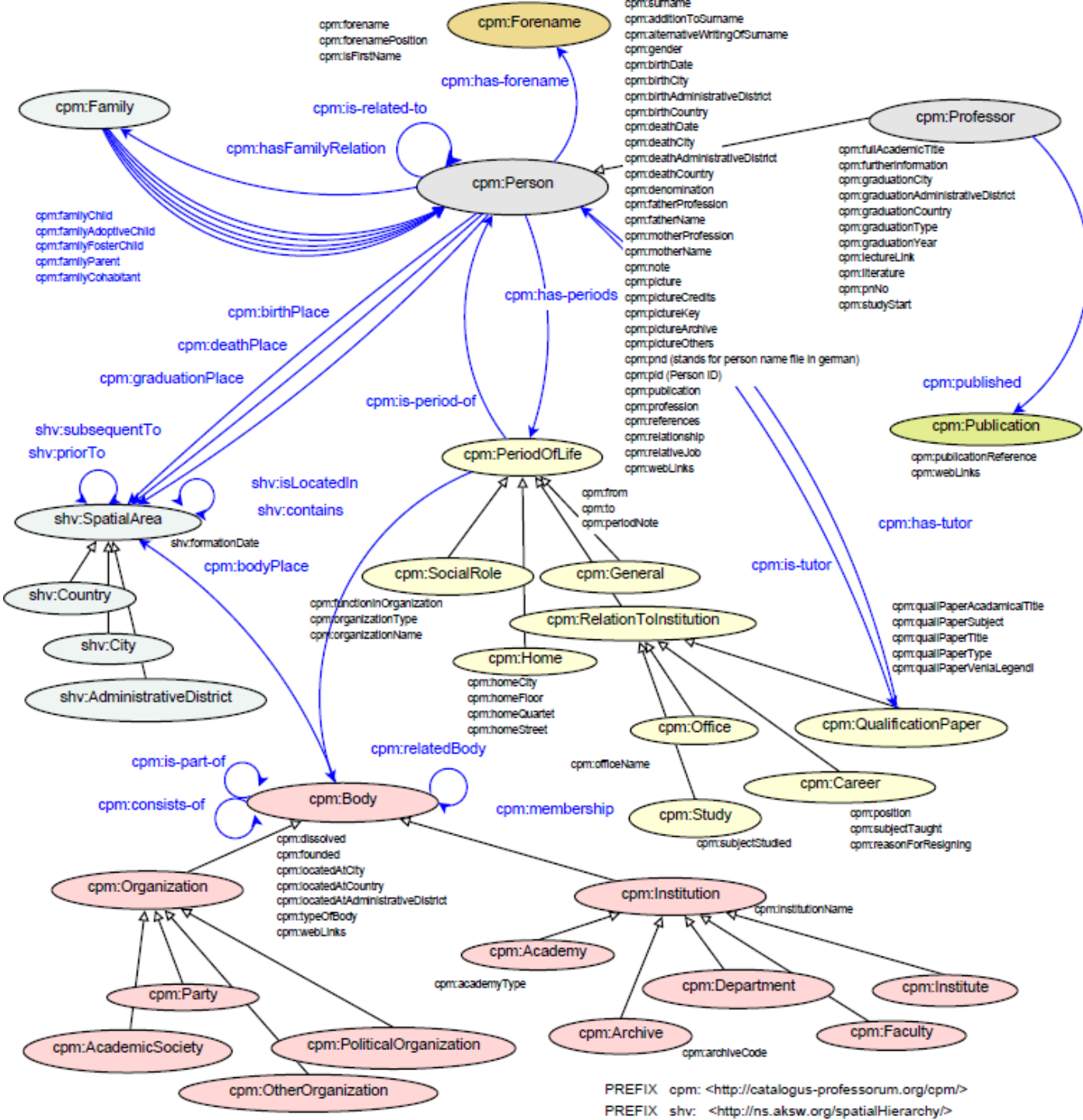
Lorenz Morsbach

Show in Main Tab »

Auftreten als Wert

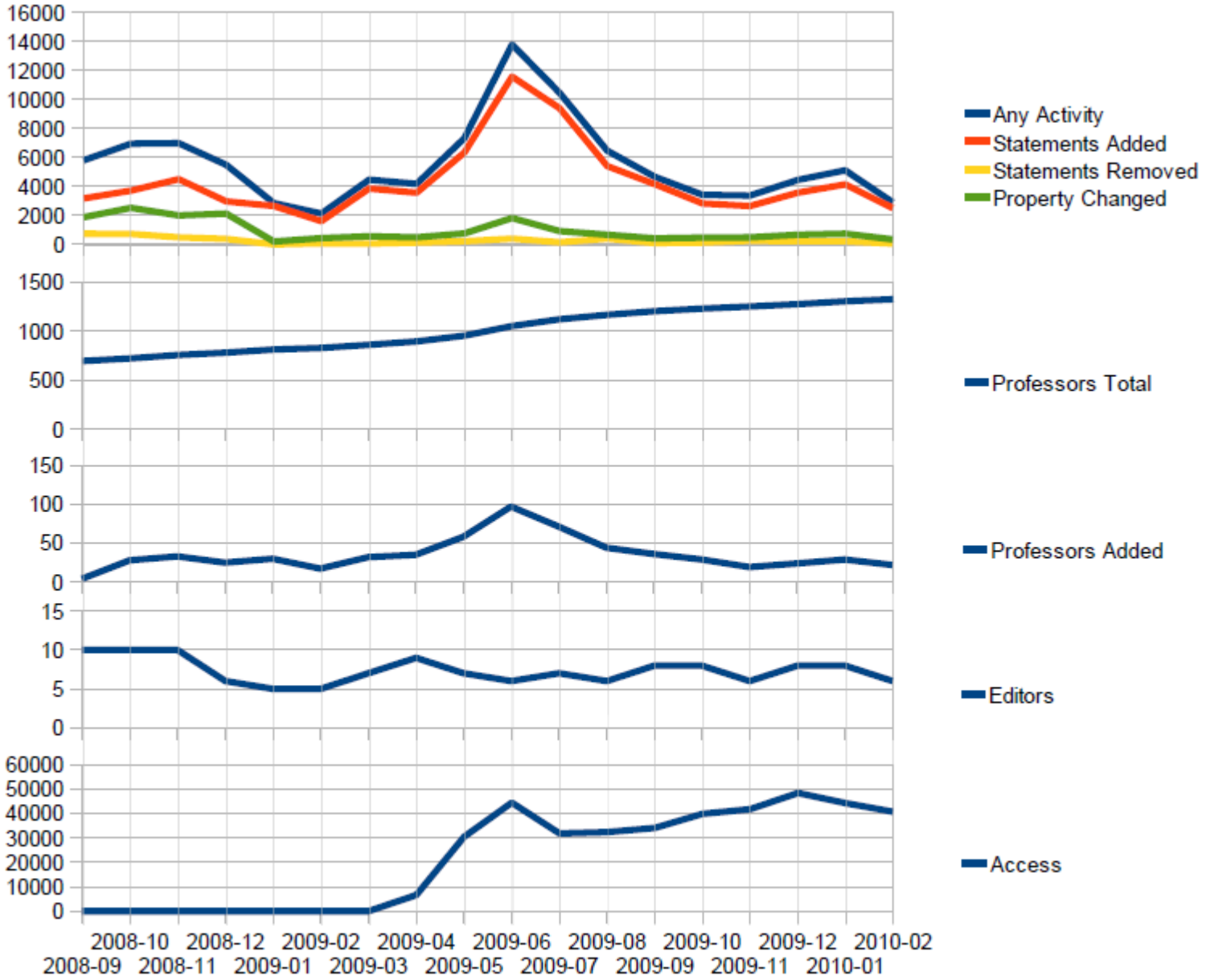
Keine Entsprechungen.

The CPL Model



LOD2
Creating Knowledge out of Interlinked Data

CPL Authoring Activity



Levin Ludwig Schücking

Geburtsstadt	Burgsteinfurt
Geburtstag	1878-05-29
Geburtsort	http://catalogus-professorum.org/place/Burgsteinfurt
birthState	Preußen
Todesstadt	Farchant
Todestag	1964-10-12
Sterbeort	http://catalogus-professorum.org/place/Farchant
voller akademischer Titel	Prof. Dr. phil.
weiter Informationen	1946-1952 Lehrstuhlvertretung an der Universität Erlangen
Abitur Stadt	Münster
Ort des Abiturs	http://catalogus-professorum.org/place/Burgsteinfurt
graduationState	Preußen
eine Person hat Vornamen	http://catalogus-professorum.org/lipsiensis/Levin-1-1 http://catalogus-professorum.org/lipsiensis/Ludwig-2-0
Person hat Lebensabschnitte	http://catalogus-professorum.org/lipsiensis/1904HabilitationinEnglischerPhilologieanderUniversittGttingen86 1925-1944 oProf für Anglistik an der Philologisch-Historischen Abteilung der Philosophischen Fakultät der Universität Leipzig http://catalogus-professorum.org/lipsiensis/Karriere_594 1910-1916 aoProf für Englische Sprache und Literatur an der Universität Jena 1916-1925 oProf für Englische Sprache und Literatur an der Universität Breslau 1952 oProf für Englische Sprache und Literatur an der Universität München Mitglied der American Academy of Arts and Sciences

1897-1901 Studium: Neuere Sprachen und Kunstgeschichte in Freiburg, Göttingen, Berlin und München

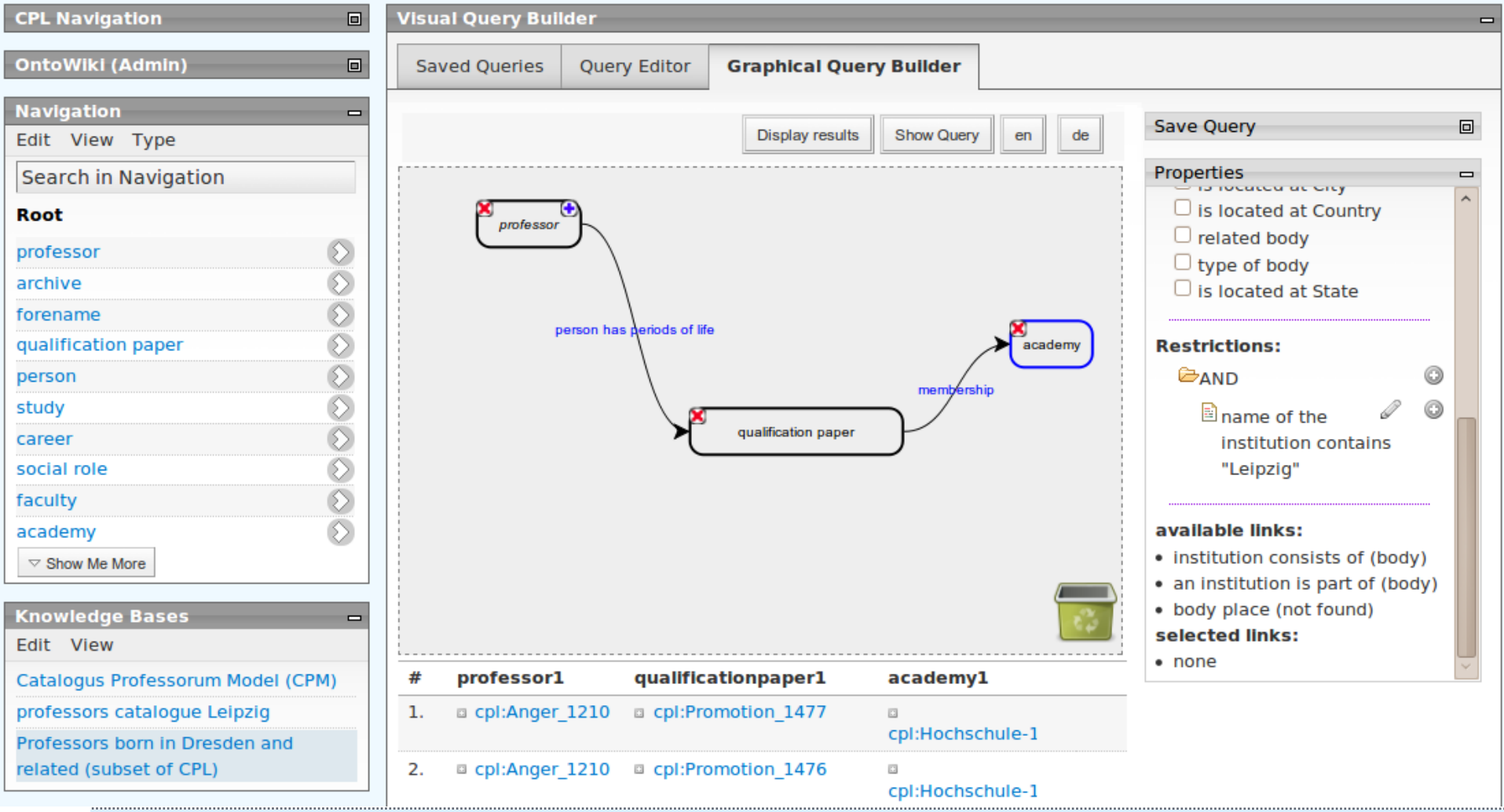
von	1897
Mitgliedschaft	Universität München
Fach	Kunstgeschichte
bis	1901-01-01
mentions	Studium Data Creation Data Item Data Creating Service
request	Request for <http://catalogus-professorum.org/lipsiensis/Study_110>
type	Studium
label	1897-1901 Studium: Neuere Sprachen und Kunstgeschichte in Freiburg, Göttingen, Berlin und München
is-period-of	Levin Ludwig Schücking
requested	Studium Fach
is used by	This Session

Find all

Table | Map | Calendar | Timeline | SPARQL | New...

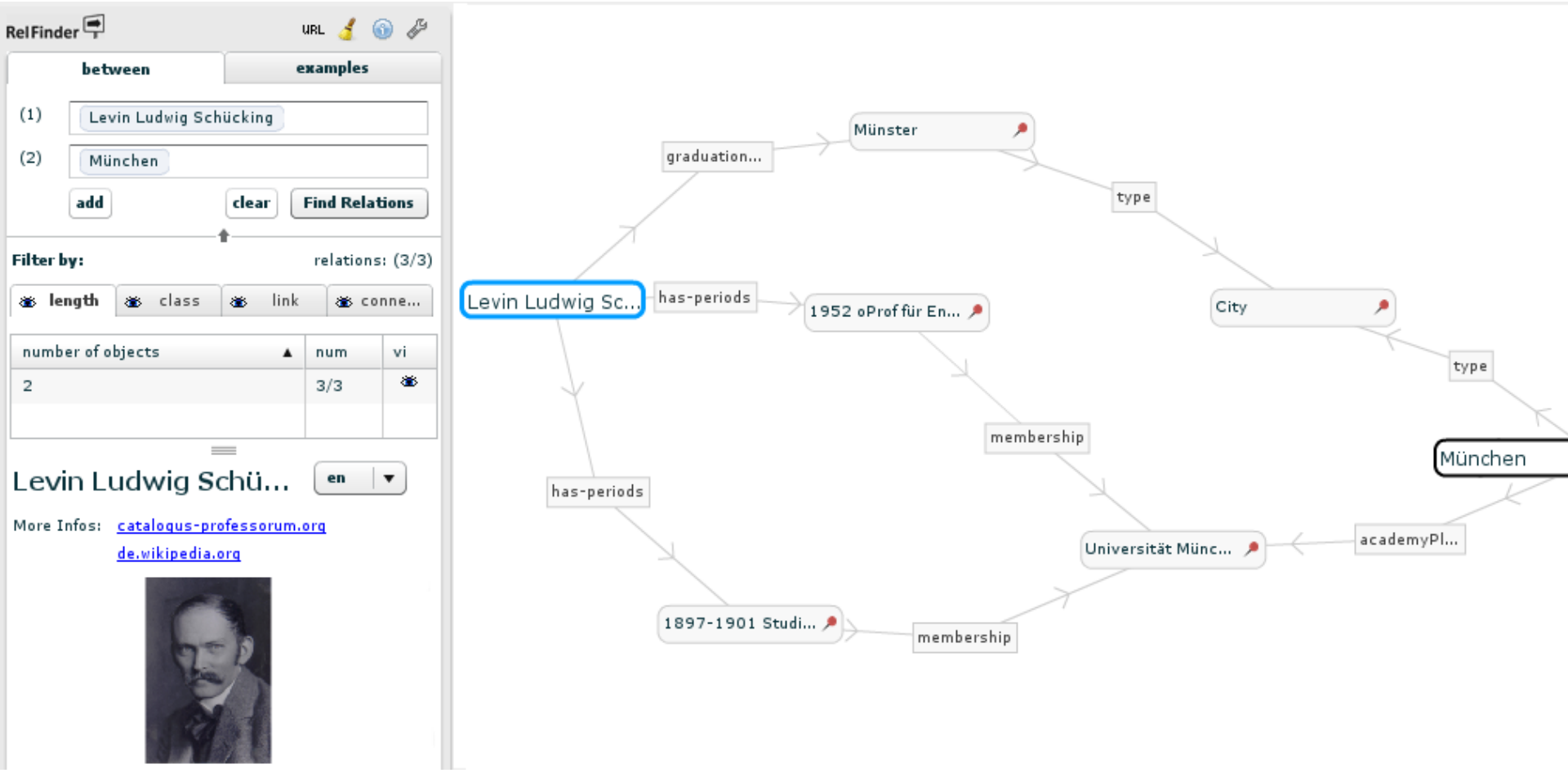
Select queries to display:

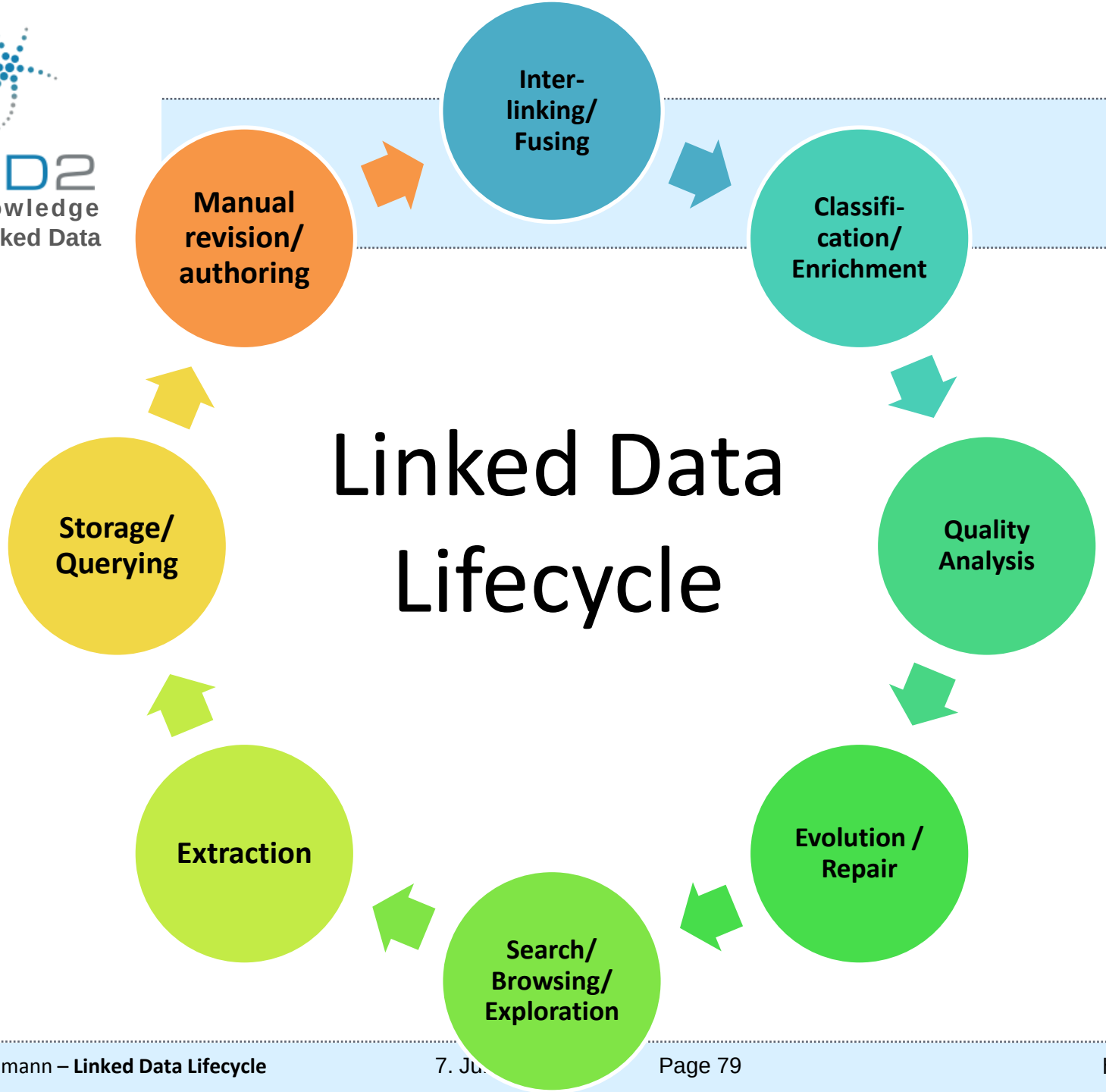
Visual Query Builder



LOD2
Creating Knowledge
out of Interlinked Data

Relationship Finder in CPL




Linked Data Lifecycle
Inter-linking/ Fusing
Classifi-cation/ Enrichment
Quality Analysis
Evolution / Repair
Search/ Browsing/ Exploration
Extraction
Storage/ Querying
Manual revision/ authoring

Make the Web a Linked Data Washing Machine

Make the Web a Linked Data Washing Machine

AKSW Data Web Building Blocks

Creating Knowledge
ou

RDB2RDF
Mapping relational data to RDF

DL-Learner
Machine Learning for Ontologies

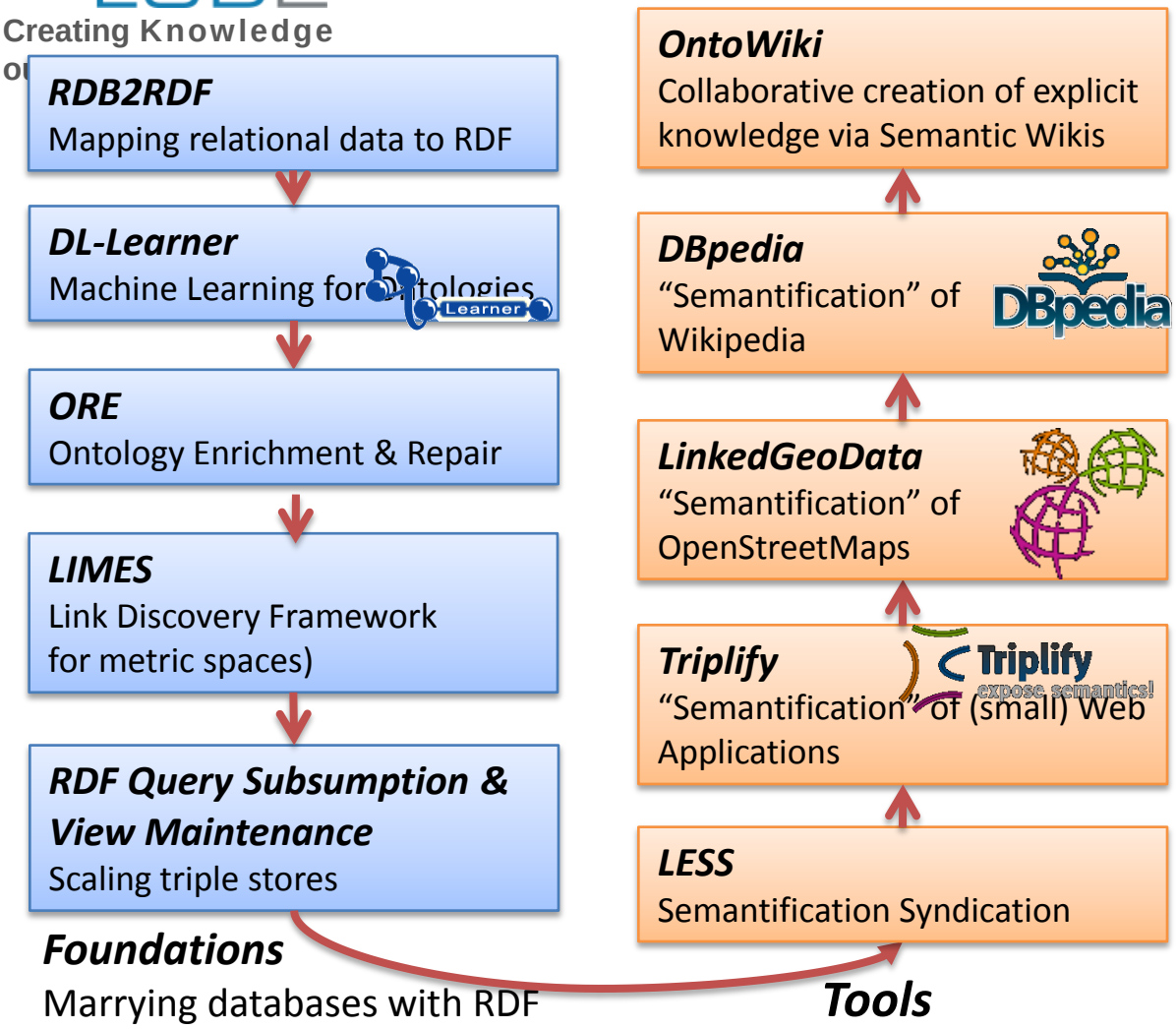
ORE
Ontology Enrichment & Repair

LIMES
Link Discovery Framework
for metric spaces)

RDF Query Subsumption & View Maintenance
Scaling triple stores

Foundations
Marrying databases with RDF and ontologies

AKSW Data Web Building Blocks



AKSW Data Web Building Blocks

RDB2RDF
Mapping relational data to RDF

DL-Learner
Machine Learning for Ontologies

ORE
Ontology Enrichment & Repair

LIMES
Link Discovery Framework for metric spaces)

RDF Query Subsumption & View Maintenance
Scaling triple stores

Foundations
Marrying databases with RDF and ontologies

OntoWiki
Collaborative creation of explicit knowledge via Semantic Wikis

DBpedia
"Semantification" of Wikipedia

LinkedGeoData
"Semantification" of OpenStreetMaps

Triplify
"Semantification" of (small) Web Applications

LESS
Semantification Syndication

Tools

Vakantieland
Building Data Web applications

SoftWiki
Distributed, stakeholder driven Requirements Engineering

OpenResearch.org
A semantic Wiki for the sciences

xOperator
Combining Instant Messaging with the Data Web

Catalogus Professorum
Prosopographical knowledge base

...

Applications
Bringing the Data Web to end users

PUBLINK - Linked Open Data Consultancy
http://lod2.eu/Article/Publink.html

Thanks for your attention!

Jens Lehmann
http://jens-lehmann.org | http://aksw.org | http://lod2.eu
lehmann@informatik.uni-leipzig.de

UNIVERSITÄT LEIPZIG