Agile Knowledge Engineering und Semantic Web - Aussen gut. Innen besser.

Thomas Riechert*, Christian Rau**

Universität Leipzig
–
* Agile Knowledge Engineering and Semantic Web und
** Lehrstuhl für Neuere und Neueste Geschichte

Werkstattberichte Vernetztes Wissen / Semantic Web
.hist2011, 14.09.2011, Berlin

Agenda

1. Über die AKSW Gruppe
2. LOD2
3. Professorenkatalog
4. Tools

Überblick

- Seit 2006 am Institut für Informatik und An-Institut für Angewandte Informatik (InfAI) der Universität Leipzig
- 20 Wissenschaftler (Leitung: Sören Auer)
- Forschungsgebiete:
  - Emergent Semantics
  - Machine Learning and Ontology Engineering
  - Semantic Abstraction
- Entwicklung von Methoden, Werkzeugen und Applikationen für adaptives Knowledge Engineering
- Grundlagenforschung Semantic Web Technologien
- Social Semantic Web
- `http://aksw.org`

Geförderte Projekte

- LOD2 – Creating Knowledge out of Interlinked Data
- LATC – LOD Around-the-Clock
- SCMS – Semantic Content Management Systems
- Social Semantic Collaboration for EKM, E-Learning E-Tourism
- LE4SW Regional Technology Platform for Social Semantic Collaboration
- SoftWiki – Semantics- Community-Based Requirements Engineering

Open Source und Community Projekte

- LinkedGeoData
- OntoWiki
- Triplify
- Cofundos
- DBpedia
- DL-Learner
- Border Flow
- LIMES
- LinkedHistory
- FOX
- ALOE
- Mobile Social Semantic Web Client (MSSW)
- ReDD Observatory
- LESS – Syndicate Linked Data Content
- Catalogus Professorum Lipsiensis
- Erfurt – PHP Semantic Web API
- Adaptive Caching SPARQL Proxy
- xOperator – Semantic Instant Messaging
- od@fmi – Open Data for the University of Leipzig's Math and Computer Science Faculty
- Open Research.org

Linked Open Data - Heute

- Erweiterung des Web mit Daten (25 Billionen Fakten)
- Lebendige weltweite Forschungscommunity
- Erste Industrieanwendungen (z.B. BBC, Thomson Reuters, Good Realations)
- Zunehmender Einsatz bei der Publikation öffentlicher Daten im E-Government
- Linked Data hat sich zur Publikation von Daten im Semantic Web etabliert

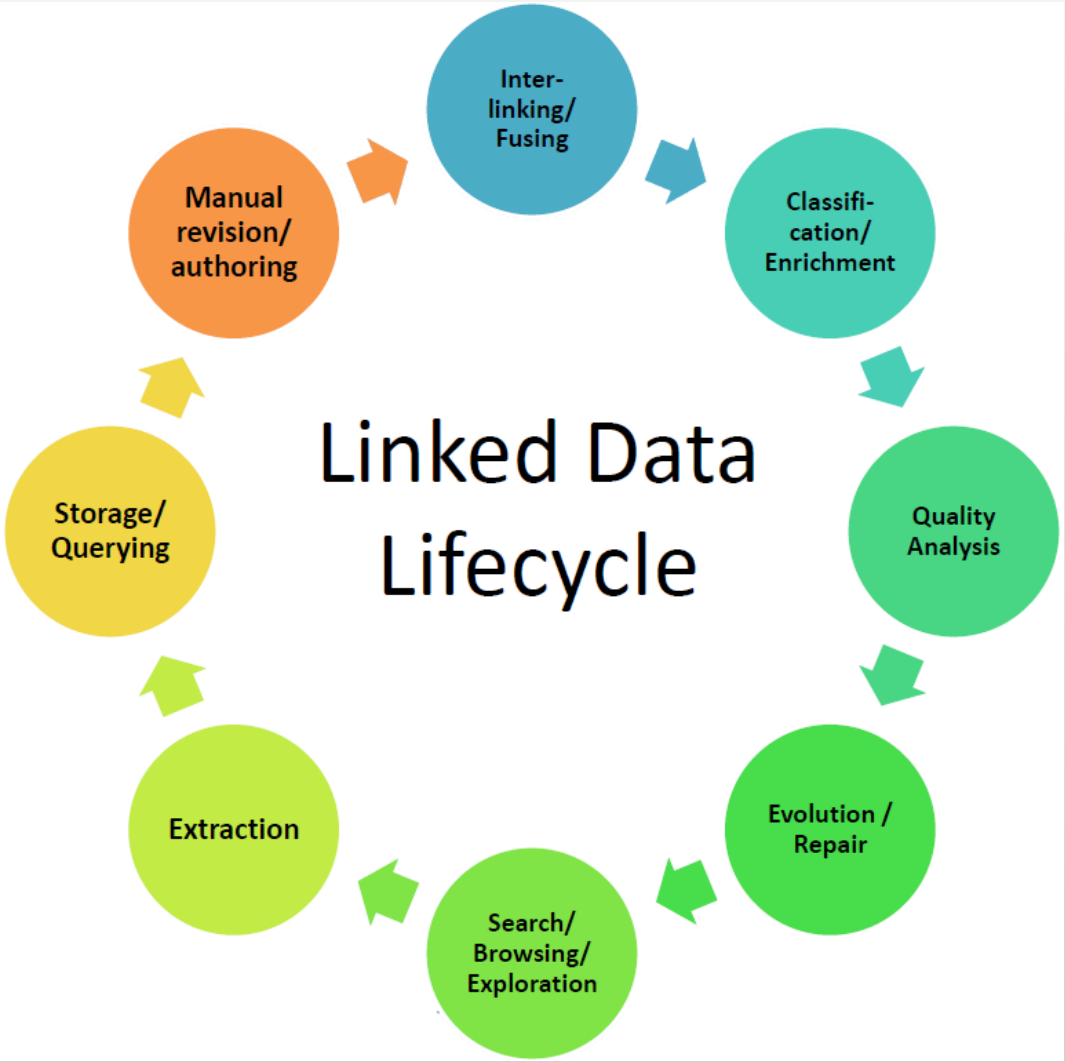
Linked Open Data - Herausforderungen

- Kohärenz: Relativ wenige, aufwändig gepflegte Links
- Qualität: Teilweise schlechte Qualität von Daten und Inkonsistenzen
- Performance: Noch erhebliche Nachteile im Vergleich zu Relationalen Datenbanken
- Daten Nutzung: Großtechnischen Verarbeitung, Schema-Mapping und Daten-Fusion noch in den Kinderschuhen
- Usability: Fehlende direkte Endbenutzer-Werkzeuge und Netzwerk-Effekte

Diese Themen sind eng miteinander verwandt und dürften letztendlich zu einem Ökosystem von verlinkten Wissen führen!

Linked Data - Lifecycle

- Extraktion
- Speicherung und Anfragen
- Erstellen und Editieren
- Verknüpfung
- Anreicherung
- Validierung
- Evolution
- Suche und Navigation



Anwendungsfall - Catalogus Professorum Lipsiensis

Hintergrund:

- 600-jähriges Jubiläum 2009
- Erste Version des Katalogs war eine einzelne Datenbank-Tabelle und erreichte die Begrenzung von 255 Spalten.
- Hoher Aufwand beim Verteilen und Synchronisieren der Tabelle
- Anforderung an die neue Datenbank:
  - Kollaboration während der Datenerhebung
  - Online Publikation
  - Begrenzte Ressourcen

Anwendungsfall - Catalogus Professorum Lipsiensis

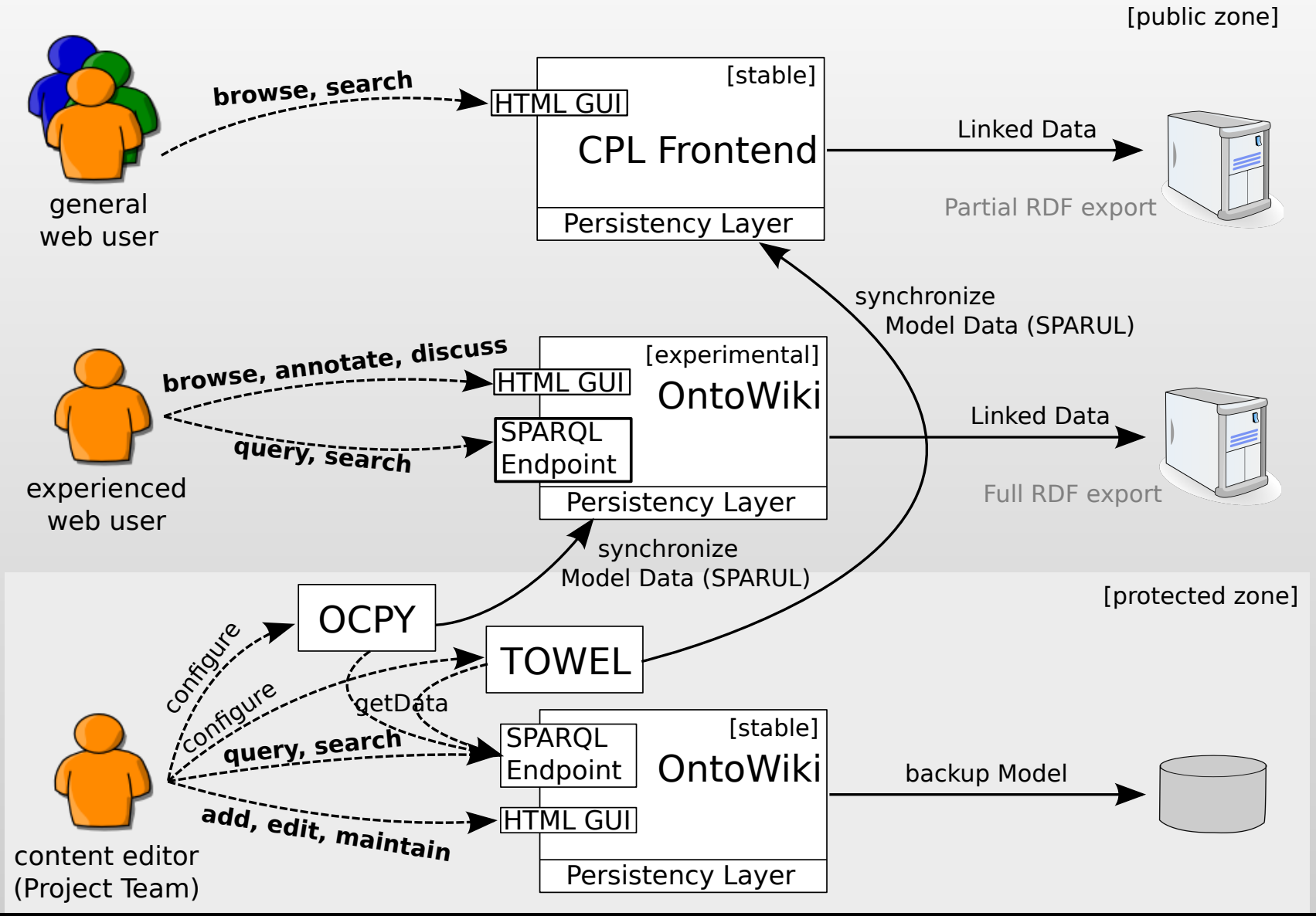
Fakten:

- 1,300 Professoren
- 10,000 verbundene Lebensabschnitte
- 400 Institutionen
- Die resultierende OWL-Wissensbasis besteht aus:
  - 200,000 Trippel im Kern des CPL
  - 173,000 Manuell hinzugefügt

Engineering Co-design-Methodologie

1. Analyse bestehender Informationen
2. Initialisierung (September 2008)
3. Wiki-basierte Datenerhebung und Ontologie-Verfeinerung
4. Publikation im Netz (Mai 2009)
5. Verlinkung mit anderen Datenbanken
6. Alignment zu anderen Ontologien (Aktuell)

CPL Architektur



OntoWiki - Überblick

- Intuitives Erstellen semantischer Inhalte,
- kollaborativ und räumlich verteilt
- Änderungsverfolgung
- Kommentierung und Diskussion
- Visuelle Repräsentation
- Verschiedene Sichten auf die Daten

OntoWiki - Screenshot

CPL Navigation

OntoWiki

Extras Help Debug

Search for Resources

POWERED BY VIRTUOSO

Knowledge Bases

View

Catalogus Professorum Lipsiensis

Catalogus Professorum Model (CPM)

Navigation: Classes

View Type

Search in Navigation

Year
person
family
body
forename
period of life
publication
SpatialThing
cpm:Brother

Login

Local OpenID FOAF+SSL

Username

Password

Remember me

Catalogus Professorum Lipsiensis

Resource

Properties Map History Community Source

Catalogus Professorum Lipsiensis

type	professor
Label	Levin Ludwig Schücking
death country	Bundesrepublik Deutschland
birth place	Königreich Preußen (Provinz Westfalen) Burgsteinfurt
graduation place	Preußen Muenster
death place	Bundesrepublik Deutschland Farchant
birthcity	Burgsteinfurt
birthday, birth date	1878-05-29
death city	Farchant
death date	1964-10-12
full academical title	Prof. Dr. phil.
graduation city	Münster
picture	
person id	144
PND	117124931

Similar Instances

professor
Gerhard Wolfgang Seeliger
Wolfgang Carl Wilhelm Ostwald
Richard Altmann
Hermann Altrock
Wilhelm Adolf Becker ...

Instances Linking Here

Is tutor^{-1}
Lorenz Morsbach

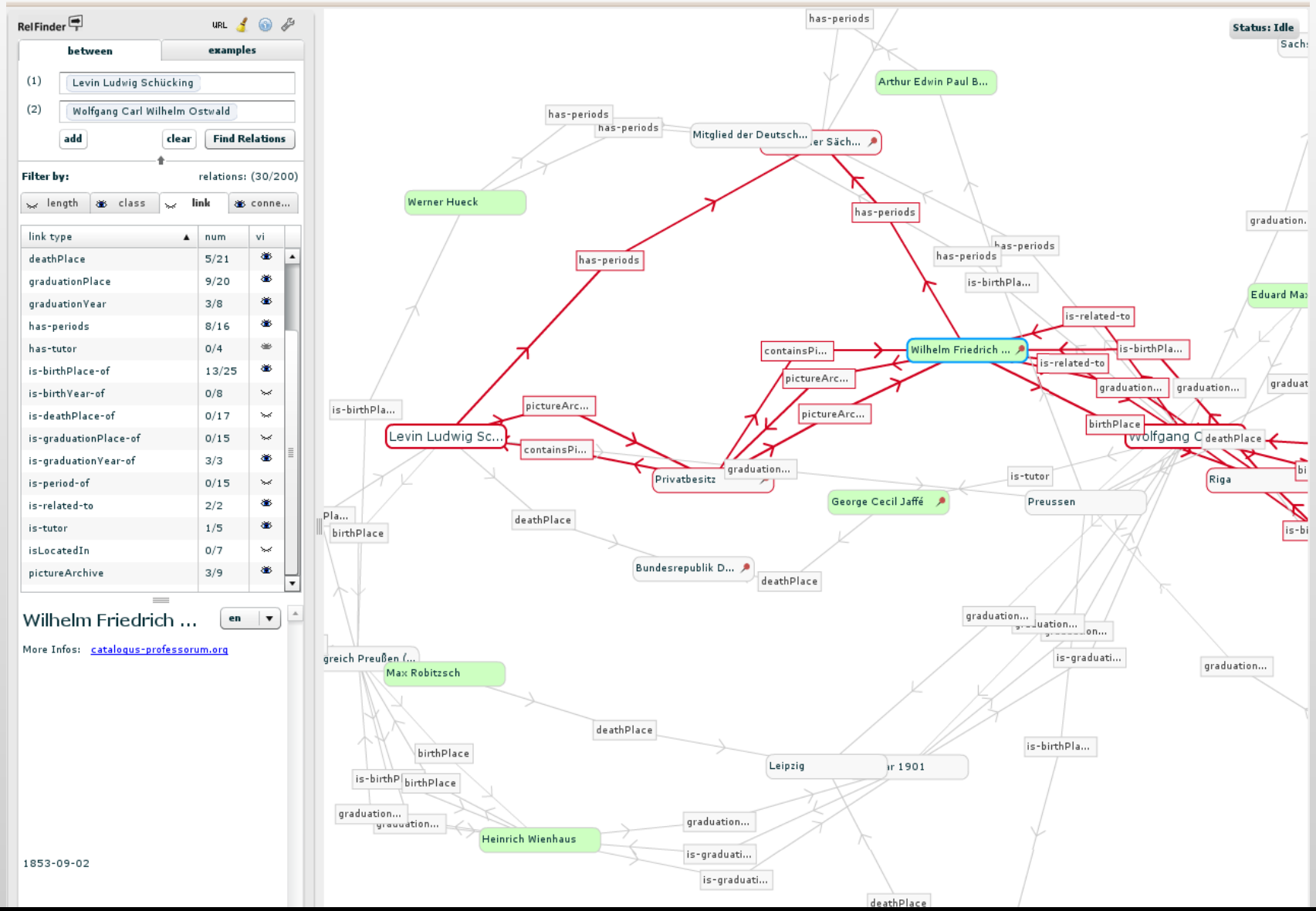
cpm:Is-birthPlace-of^{-1}
Königreich Preußen (Provinz Westfalen)
Burgsteinfurt

cpm:Is-birthYear-of^{-1}
The year 1878

cpm:Is-deathPlace-of^{-1}
Bundesrepublik Deutschland
Farchant

cpm:Is-deathYear-of^{-1}
The year 1964

Relationship-Finder



Vielen Dank für Ihre Aufmerksamkeit.

- http://aksw.org
- http://catalogus-professorum.org
- http://www.uni-leipzig.de/unigeschichte/professorenkatalog

- Thomas Riechert
- Universität Leipzig
- riechert@informatik.uni-leipzig.de