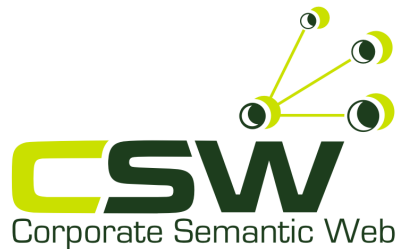


# DBpedia Deutschland

<http://de.dbpedia.org/>

Berlin, 15. September 2009  
Mario Rothe



Arbeitsgruppe  
**Corporate Semantic Web**  
Freie Universität Berlin



# Allgemeines

- Einsatz des deutschen Teiles der DBpedia
- semantische Informationen der deutschen Wikipedia
- momentan 27,826,688 Fakten
  - statischer deutscher DBpedia Dump (Version 3.3)
  - Nutzen der Live Extraction in Planung
- Ziel: Linked Data Verbund zwischen den nationalen DBpedia Datensätzen
  - ermöglicht über Interlanguage Links
- Alles noch Beta!

# Technik

## ■ AllegroGraph RDFStore

- effiziente persistente Speicherung von semantischen Daten
- Einschub: RDF (Resource Description Framework)
  - Subjekt – Prädikat – Objekt Triple
  - Berlin Anzahl der Einwohner 3.431.675
- spezielles Feature: räumliches und zeitliches Reasoning
  - erlaubt (zukünftig) Abfragen der Form:  
„Welche wichtigen historischen Ereignisse spielten sich im 18. Jahrhundert im Umkreis von 100 km von Paris ab“

# Zugriffsmöglichkeiten

## ■ SPARQL Endpoint

- <http://de.dbpedia.org/sparql>
- Webservice für Software-Klienten

## ■ AllegroGraph WebView

- SPARQL Interface um interaktiv Daten abzufragen
- erreichbar über [de.dbpedia.org](http://de.dbpedia.org)

# SPARQL

- Abfragesprache für semantische gelagerte Daten
  - standardisiert durch W3C:
    - <http://www.w3.org/TR/rdf-sparql-query/>
  - ähnlich wie SQL für Datenbanken:

```
SELECT <return variables>
WHERE {
    <triple pattern>
}
```

- Pattern: ?name rdf:type foaf:Person

# SPARQL Beispiel

„... gebe mir die Kurzinformationen aller Personen, die am 8. Januar 1923 geboren wurden“

## ■ Als SPARQL:

```
SELECT ?name ?abstract
WHERE {
  ?name rdf:type foaf:Person ;
        dbproperty:abstract ?abstract ;
        dbproperty:birth "1923-01-08"^^xmls:date .
}
```

# Anwendungsmöglichkeiten

# Anwendungen

## ■ Wikipedia

- wohl größte lexikalische Ansammlung von Informationen aller Zeiten
- User Generated Content
  - Aktualität hervorragend
  - Qualität ausreichend
- frei verfügbar!

## ■ DBpedia:

- maschineninterpretierbare Kontext dieser Informationen
- Informationskombination und Reasoning



# Anwendungsmöglichkeiten

## ■ Informationsaggregation

- Informationsportale aller Art
  - Städte, Museen, Touristeninformationen ...
- Nachrichtenportale
  - weiterführende Informationen zum Nachrichteninhalt
- fachspezifische Foren
  - Erklärung von Fachbegriffen
  - wie Google AdWords
- Unternehmen
  - fachübergreifendes Hintergrundwissen zu Firmeninformationen (im CMS)

## ■ Reasoning auf DBpedia Daten

- „Welche Beziehungen existieren zwischen Albert Einstein und Kurt Gödel