DBpedia
A crystallization point for the Web of Data

Christian Bizer, Jens Lehmann, Georgi Kobilarov, Soren Auer, Christian Becker, Richard Cyganiak, Sebastian Hellmann

Presented by Jeremy Chen

Motivation

- Most existing knowledge bases cover only specific domains
- Costly to keep existing knowledge bases up-to-date as domains change

At the same time…..

- Wikipedia is a big central knowledge source
- **DBpedia leverages this gigantic source of knowledge by extracting structured information**

Advantages

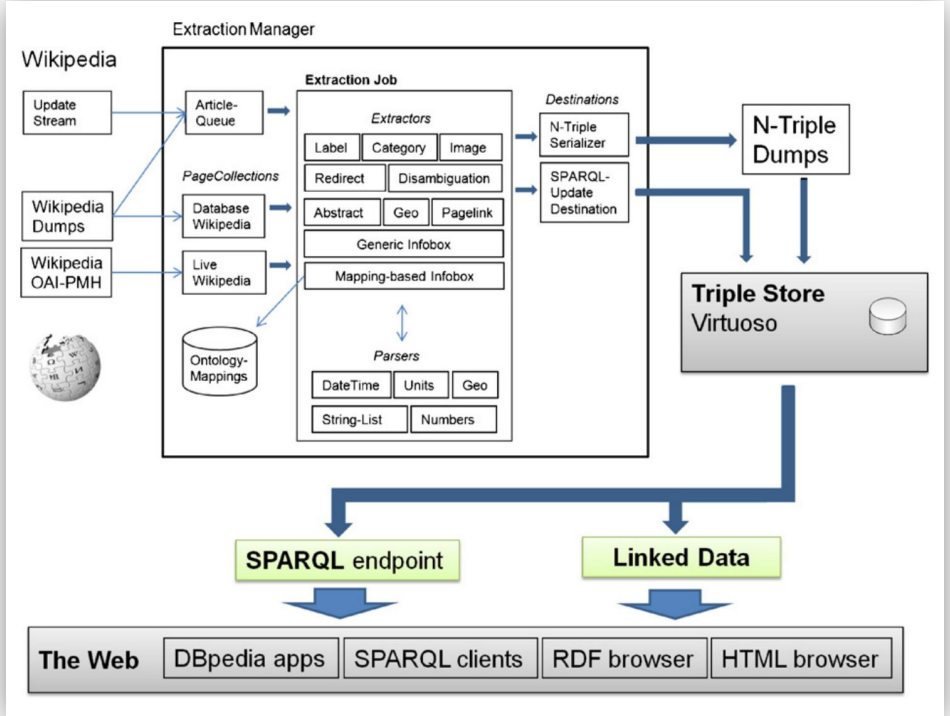
- Covers many domains
- Evolves automatically as Wikipedia changes
- Truly multilingual
- Accessible on the Web

Outline

- DBpedia knowledge extraction framework
- DBpedia knowledge base
- Different access mechanisms
- Interlinked Web content with DBpedia
- Applications

DBpedia Knowledge Extraction Framework

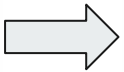
- Workflows
  - Dump-based extraction
    - Monthly SQL dumps
  - Live extraction



Reference: DBpedia - A crystallization point for the Web of Data

DBpedia Knowledge Extraction Framework

- Infobox Extraction
  - Generic infobox extraction
    - *Subject*: DBpedia URI
    - *Predicate*: http://dbpedia.org/property/ + attribute name
    - *Object*: attribute value
  - Mapping-based infobox extraction



{{ Infobox Actor

| birthname = Thomas Jeffrey Hanks

| birthdate = {{birth date and age|1956|7|9}}

| birthplace = [[Concord, California|Concord]],
[[California]]

| yearsactive = 1979 - present

| occupation = Actor, producer, director,
[[voice over artist]],
writer, speaker

Fig. 2. Infobox Tom Hanks.

Reference: DBpedia - A crystallization point for the Web of Data

DBpedia Knowledge Base

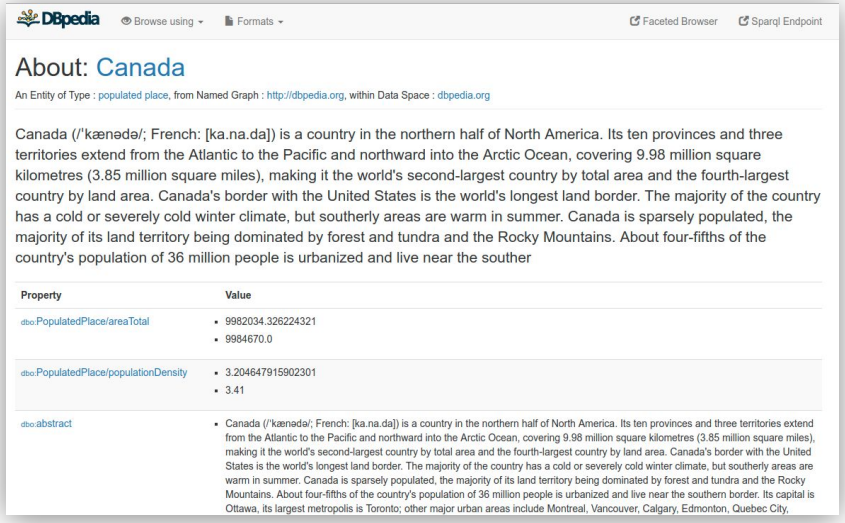
- Entity Identifier: English article name
- Entity Properties
  - General Properties
    - Label
    - English abstract
    - Links to external Web pages
    - etc.
  - Infobox-specific properties

Accessing DBpedia

- 4 access mechanisms
  - Linked Data
  - SPARQL endpoint
  - RDF dumps
  - Lookup index

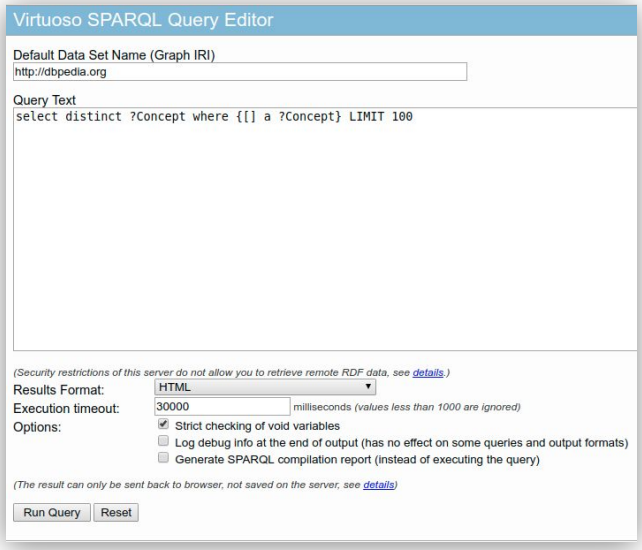
Accessing DBpedia - Linked Data

- *Semantic Web agents*: RDF descriptions
- *Traditional Web browsers*: simple HTML view



Accessing DBpedia - SPARQL Endpoint

- http://dbpedia.org/sparql



Concept
http://www.w3.org/2002/07/owl#FunctionalProperty
http://www.w3.org/1999/02/22-rdf-syntax-ns#Property
http://www.w3.org/2002/07/owl#Thing
http://www.w3.org/2002/07/owl#Class
http://www.w3.org/2002/07/owl#Ontology
http://www.w3.org/2002/07/owl#ObjectProperty
http://www.w3.org/2002/07/owl#DatatypeProperty
http://xmlns.com/foaf/0.1/Organization
http://dbpedia.org/ontology/Company
http://xmlns.com/foaf/0.1/Person
http://dbpedia.org/ontology/Activity
http://dbpedia.org/ontology/Name
http://dbpedia.org/ontology/Person
http://dbpedia.org/ontology/Actor
http://dbpedia.org/ontology/Place
http://dbpedia.org/ontology/Publisher
http://dbpedia.org/ontology/Genre
http://dbpedia.org/ontology/Language
http://dbpedia.org/ontology/Location
http://dbpedia.org/ontology/Software
http://dbpedia.org/ontology/School
http://xmlns.com/foaf/0.1/Document
http://purl.org/ontology/bibo/Book
http://purl.org/vocommons/voaf#Vocabulary

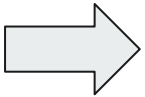
Accessing DBpedia - RDF Dumps

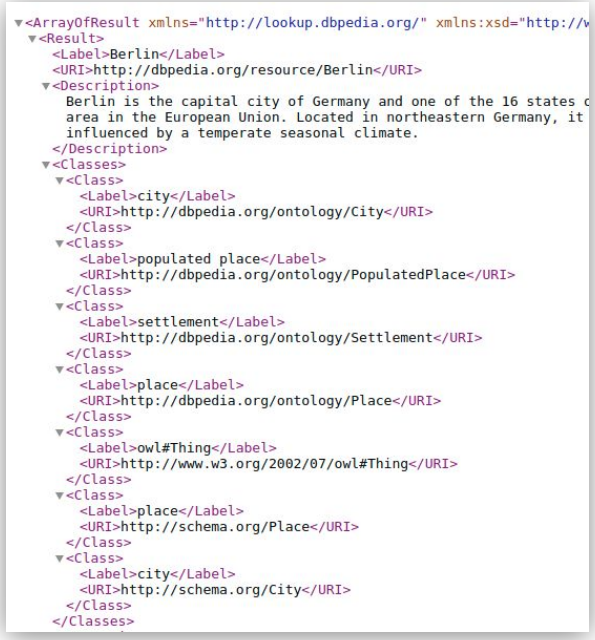
- Sliced DBpedia knowledge base by triple predicate
  - e.g. https://wiki.dbpedia.org/dbpedia-version-2016-04

Accessing DBpedia - Lookup Index

- Label -> DBpedia URI

http://lookup.dbpedia.org/api/search/KeywordSearch?QueryClass=place&QueryString=berlin





```
<ArrayOfResult xmlns="http://lookup.dbpedia.org/" xmlns:xsd="http://w
  <Result>
    <Label>Berlin</Label>
    <URI>http://dbpedia.org/resource/Berlin</URI>
    <Description>
      Berlin is the capital city of Germany and one of the 16 states o
      area in the European Union. Located in northeastern Germany, it
      influenced by a temperate seasonal climate.
    </Description>
    <Classes>
      <Class>
        <Label>city</Label>
        <URI>http://dbpedia.org/ontology/City</URI>
      </Class>
      <Class>
        <Label>populated place</Label>
        <URI>http://dbpedia.org/ontology/PopulatedPlace</URI>
      </Class>
      <Class>
        <Label>settlement</Label>
        <URI>http://dbpedia.org/ontology/Settlement</URI>
      </Class>
      <Class>
        <Label>place</Label>
        <URI>http://dbpedia.org/ontology/Place</URI>
      </Class>
      <Class>
        <Label>owl#Thing</Label>
        <URI>http://www.w3.org/2002/07/owl#Thing</URI>
      </Class>
      <Class>
        <Label>place</Label>
        <URI>http://schema.org/Place</URI>
      </Class>
      <Class>
        <Label>city</Label>
        <URI>http://schema.org/City</URI>
      </Class>
    </Classes>
```

Interlinked Web Content

- More data publishers set RDF links to DBpedia entities
- Together with outgoing edges published by DBpedia, DBpedia becomes a central interlinking hub of Web of Data
  - Web of Data Browsing and Crawling
    - Linked Data browser, Semantic Web search engines
  - Web Data Fusion and Mashups
    - Fusing data from different sources to generate integrated views
  - Web Content Annotation

Interlinked Web Content - Example

- Complementary data about Spain
  - Agents can follow these links to retrieve additional information

<http://dbpedia.org/resource/Spain> owl:sameAs

http://www4.wiwiss.fu-berlin.de/eurostat/resource/countries/Espa%C3%B1a ;

http://rdf.freebase.com/ns/guid.9202a8c04000641f8000000000034e30 ;

http://www4.wiwiss.fu-berlin.de/factbook/resource/Spain ;

http://sw.opencyc.org/2008/06/10/concept/Mx4rvVjowpwpEbGdrcN5Y29ycA .

- Annote a research paper with “Data Integration”

<http://data.semanticweb.org/conference/eswc/2008/paper/356>

swc:hasTopic <http://dbpedia.org/resource/Data_integration> .

Reference: DBpedia - A crystallization point for the Web of Data

Applications

- Browsing and exploration
  - DBpedia Mobile
- Querying and search
  - DBpedia Query Builder
  - Relationship Finder
- Content annotation

Applications - DBpedia Mobile

- Location-aware client for the Semantic Web

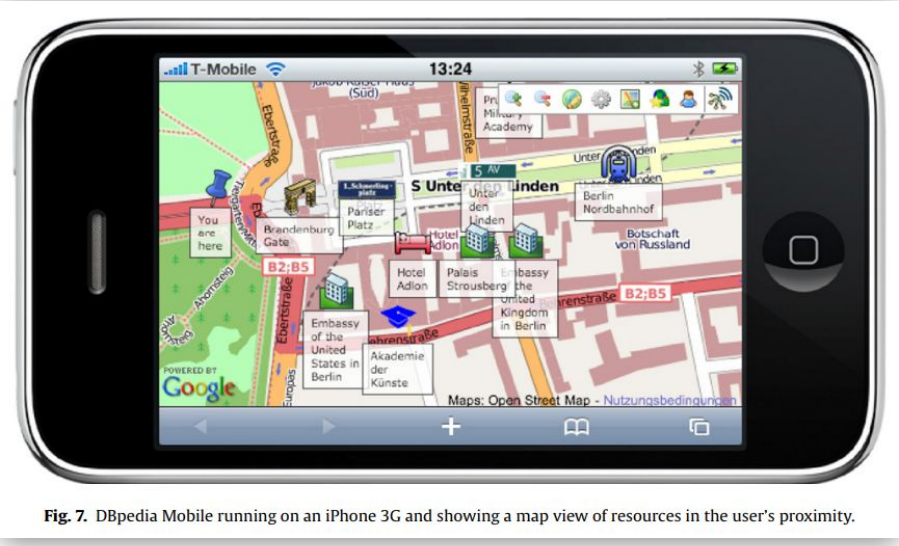


Fig. 7. DBpedia Mobile running on an iPhone 3G and showing a map view of resources in the user's proximity.

Reference: DBpedia - A crystallization point for the Web of Data

Applications - DBpedia Query Builder

- Users can hardly know the properties/identifiers in the knowledge base

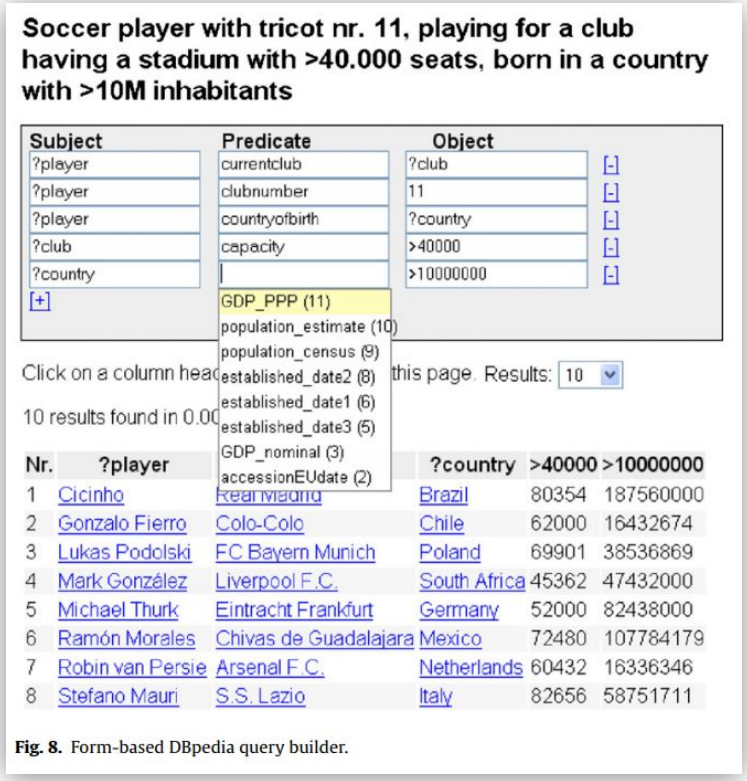


Fig. 8. Form-based DBpedia query builder.

Reference: DBpedia - A crystallization point for the Web of Data

Applications - Relationship Finder

- Find connections between two different entities in DBpedia

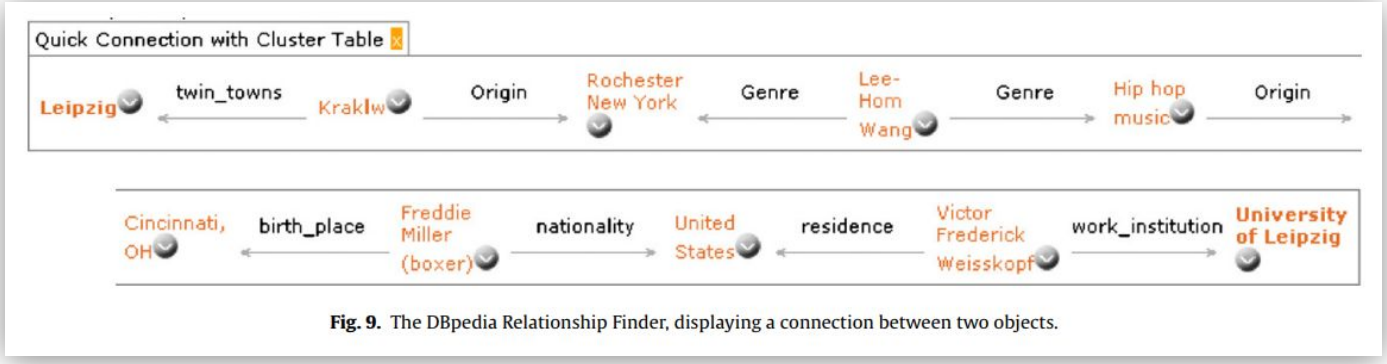


Fig. 9. The DBpedia Relationship Finder, displaying a connection between two objects.

Reference: DBpedia - A crystallization point for the Web of Data

Thank you!