DBpedia
A crystallization point for the Web of Data


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

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
Accessing DBpedia - RDF Dumps

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

DBpedia version 2016-04

Dataset category: DBpedia release
Publication Year: 2016
This release is based on updated Wikipedia dumps dating from March/April 2016 featuring a significantly expanded base of information as well as richer and (hopefully) cleaner data based on the DBpedia ontology.
Accessing DBpedia - Lookup Index

- Label -> DBpedia URI

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

- Annote a research paper with “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

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

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!