The Linked Open Data Wiki
Empowering Organizational Knowledge Bases with Linked Open Data

SMWCon Fall 2017, Rotterdam, NL

Matthias Frank, Research Scientist

05. Oktober 2017
Motivation

- Easy reuse of public available information for organizational knowledge bases

**Example:** Knowledge base for geo a marketing campaign in SMB

- *Public information:* population, coordinate location, area, elevation, sister cities, time zone, postal codes, coat of arms, local dialling code, licence plate code, …

- *Organizational knowledge:* customers, stores, employees, market survey, competition analysis, …
Levels of Open Data

Berners-Lee (2006)
12 Datasets
25 Datasets
28 Datasets
28 Datasets
45 Datasets

As of September 2008
89 Datasets
93 Datasets
203 Datasets
295 Datasets
### Semantic Wiki Software (Open Source)

<table>
<thead>
<tr>
<th>Approach</th>
<th>Underlying Engine</th>
<th>Data Storage / Export</th>
<th>Usage</th>
<th>Query construction</th>
<th>Integration of LOV/LOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>MW / Cargo</td>
<td>MediaWiki</td>
<td>Relational, CSV export</td>
<td>Wikipedia, SMBs</td>
<td>#cargo_query (SQL-like)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantic MW</td>
<td>MediaWiki</td>
<td>Relational, RDF mirror / export</td>
<td>Organizational knowledge management, e.g. Organizational knowledge graphs</td>
<td>#ask: (SPARQL)</td>
<td>manual import of single terms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OntoWiki</td>
<td>-</td>
<td>Relational or RDF</td>
<td>Organizational knowledge graphs</td>
<td>SPARQL</td>
<td>publish ontology with LOV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Research Questions

- **RQ1**: How can we assist users of organizational wikis in establishing meaningful (i.e., semantic) relations to Linked Open Data entities?
- **RQ2**: How can we keep track of the provenance of statements in an organizational wiki, especially if these statements are inferred or gathered from Linked Open Data?
- **RQ3**: How can we evaluate and interpret potential uncertain, incomplete, inconsistent or redundant Linked Open Data correctly in order to increase the informative value of an organizational knowledge base?
Approach

Linked Data Wiki
Linked Data Wiki Architecture

Web knowledge uses as information base
Linked Data Wiki Architecture

Linked-Data-Management

- Load from LOD
- Publish as LOD
Linked Data Wiki Architecture

Local Information
- Cache web knowledge
- Store additional information
- Load information
- Update information
Linked Data Wiki Architecture

MW as user interface
- Load statements
- Add statements
- Update statements
Linked Data Wiki Architecture

Presentation data
- Free text
- Place holder
- MW-markup
Implementation
New Category

Formular: Kategorie

Dies ist das Formular „Kategorie“. Um eine Seite mit dieser automatisch zum Bearbeitungsformular der Seite weitergeleitet zu bekommen.

Stadt
Erstellen oder bearbeiten

SELECT * WHERE {
?category rdf:type rdfs:Class;
rdfs:label "Stadt".
} limit 100

- schema.org:City
- dbpedia-owl:City
- wikidata:Q515
New Instance

New page: “Karlsruhe”
Category: “Stadt”

SELECT * WHERE {
?instance rdf:type wikidata:Q515;
rdfs:label "Karlsruhe".
} limit 100
Schema knowledge

- Attributes of a city relevant for geo marketing:
  - population
  - coordinate location
  - area
  - elevation
  - sister cities
  - time zone
  - postal codes
  - coat of arms
  - local dialling code
  - licence plate code
Identifying incomplete data

- Attributes of “Karlsruhe”:
  - population: 296,033
  - coordinate location: 49°0'50"N, 8°24'15"E
  - area: 173.46 square kilometre
  - elevation: 115 metre
  - sister cities: Nancy, Nottingham, Krasnodar, Timișoara, Halle
  - time zone: UTC+01:00
  - postal codes: 76229, 76131, 76137, 76133, 76135, 76139, 76149, 76199, 76185, 76187, 76189, 76227, 76228
  - coat of arms: [image]
  - local dialling code: 
  - licence plate code: KA
Discussion
Research Question 1

- How can we assist users of organizational wikis in establishing meaningful (i.e., semantic) relations to Linked Open Data entities?

Our contribution:
- Query LOD automatically based on category and label
Research Question 2

How can we keep track of the provenance of statements in an organizational wiki, especially if these statements are inferred or gathered from Linked Open Data?

Our contribution:

✓ Store provenance information for each statement in the local repository using the Open Annotation Data Model
Research Question 3

How can we evaluate and interpret potential uncertain, incomplete, inconsistent or redundant Linked Open Data correctly in order to increase the informative value of an organizational knowledge base?

Ongoing Work:

- **Uncertain:**
  - Evaluate rules for data provenance
  - Compare values from different (independent) sources

- **Incomplete:**
  - Use schema knowledge from LOD to determine missing values

- **Inconsistent and redundant:**
  - Evaluate rules for data provenance
  - Use page ranks to get the most likely values
THANK YOU