BERD: The knowledge graph of German companies

Dr. Renat Shigapov (Data Scientist) & Dr. Irene Schumm (Head of Research Data Center)
Mannheim University Library, 23.02.2021
Motivation: German company data is fragmented over many providers, registers and time spans

**Modern**
- orbis
- openctrorporates

**Problems**
- 😞 No unique identifier over providers
- 😞 No unique identifier over time

**Historical**
- LEI
- Bisnode

**Consequences**
- 😞 Limited Accessibility & Researchability
- 😊 Example: Named entity linking on texts and tables

**Solution**
- 😊 BERD knowledge graph of German companies
Phase 1: BERD knowledge graph construction & testing

BERD ontology and BERD data

Python + WikidataIntegrator (a wrapper of the Wikibase API)

SPARQL request

Front-end
- Python
- Flask
- JavaScript

Back-end
- Wikibase

SPARQL response
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Deutsche Telekom</td>
<td>mobile phone industry</td>
<td>Bonn</td>
<td>1994-01-01</td>
<td><a href="https://www.telekom.com/">https://www.telekom.com/</a></td>
</tr>
<tr>
<td>2</td>
<td>Lufthansa</td>
<td>air transport</td>
<td>Cologne</td>
<td>1953-01-01</td>
<td><a href="https://www.lufthansa.com">https://www.lufthansa.com</a></td>
</tr>
<tr>
<td>3</td>
<td>SAP SE</td>
<td>software development</td>
<td>Walldorf</td>
<td>1972-04-01</td>
<td><a href="https://www.sap.com">https://www.sap.com</a></td>
</tr>
</tbody>
</table>

https://github.com/UB-Mannheim/bbw
Phase 3: Speeding up and automating knowledge graph construction with RaiseWikibase

A wrapper of the Wikibase API

😊 Create and edit 1 million entities per 2-12 days.
😊 Adding the templates, modules and texts takes time.

RaiseWikibase (a wrapper of MariaDB)

😊 Create and edit 1 million entities per hour.
😊 The templates, modules and texts can be easily and quickly uploaded.

50 times speed up!

Soon at ESWC 2021 and https://github.com/UB-Mannheim/RaiseWikibase
Takeaway 1: Wikibase is an excellent software for research data management!

Takeaway 2: Let’s create more tools that will make Wikibase even better!

Takeaway 3: BERD knowledge graph for accessibility and researchability of Business, Economic and Related Data.

https://github.com/shigapov
Renat.Shigapov@bib.uni-mannheim.de
Irene.Schumm@bib.uni-mannheim.de