Using Wikidata in Bibliographical Collections

an example with dblp.org

Florian Reitz @ LZI Schloss Dagstuhl, Germany

What is dblp

dblp indexes computer science and related fields.
4.7 million papers / 2.3 million author profiles
We started to integrate Wikidata in our workflow in 2016. This poster shows use cases we identified so far. This is work in progress.

382k publications in dblp mapped to Wikidata entities (via DOI)
26k author profiles mapped to Wikidata entities (manually + identical ORCID)

Application: Enrich Profiles

**Problem:** Many author profiles have little or no information on the person itself.

**Solution:** Use Wikidata to provide structured data and links to Wikipedia articles.

<table>
<thead>
<tr>
<th>occupation</th>
<th>employer</th>
<th>ORCID</th>
<th>sex or gender</th>
<th>educated at</th>
<th>award received</th>
<th>date of birth</th>
</tr>
</thead>
</table>

20% of the profiles linked to Wikidata have an article in the English Wikipedia.

- Useful for author disambiguation, statistics ...
- User support + better search results

Application: Identify Missing Data

**Problem:** Important publications are often missing in dblp because a conference or a journal has been overlooked.

**Solution:** Use known mappings to Wikidata person entities and study their publications

- Derive missing conferences/journals from missing publications
- Find new events/issues of known conferences/journals
- Reduce thematic bias by considering the wider selection of Wikidata

Application: Find Defective User Profiles

**Problem:** Mapping publications with the correct user profile is difficult and time consuming. Mappings are often incorrect and cannot be resolved algorithmically.

**Solution:** Compare our own mapping with mapping of Wikidata to find potential problems.

A potential homonym (one profile multiple persons) in dblp or a synonym (one person, multiple profiles) in Wikidata.

Clusters of suspicious mappings can help to identify problematic data sections.

Resolving conflicts can help dblp and Wikidata

Reuse the manual work the other project has already invested.