

ABSTRACT

Wikidata is an open knowledge graph platform with over 100 million entities, but its academic coverage remains limited, with many co-authors, scholarly articles, and proceedings still missing. This limitation reduces its usefulness for scientific analysis. This study proposes a semi-automated system to enrich academic entities in Wikidata using structured metadata from DBLP. The system follows an ETL pipeline consisting of data extraction, rule-based disambiguation via the Wikidata Reconciliation API, and automatic entity loading using Wikibase Integrator. The disambiguation process prioritizes attributes such as DBLP ID, ORCID, affiliation, and occupation. Evaluation was conducted on nine researchers using knowledge graph coverage and correctness metrics. The results show that the system achieved 100% coverage across all tested entities and reached a disambiguation accuracy of 80.22%. The main contribution of this study is a rule-based disambiguation system that is efficient, accurate, and scalable for enriching academic entities in Wikidata.

Keywords: *knowledge graph, wikidata, dblp, disambiguation, academic entity, data integration*