CounQER: A System for Discovering and Linking Count Information in Knowledge Bases

Shrestha Ghosh, Simon Razniewski, Gerhard Weikum
Max Planck Institute for Informatics, Germany

Problem Statement

Identify predicates in KBs storing count information - through integers, *counting predicates*, and entity enumerations, *enumerating predicates*,
Semantically align such *set predicates*.

Understanding count information has benefits in

- **KB curation**. Discover incompleteness and/or inconsistencies through alignments.
- **QA enhancements.** Enable query result enrichment and debugging. Highlight variations in predicate usage for the same concept.

Method

Supervised Classification for identifying
Counting predicates:
numberOfChildren, staffSize
Enumerating predicates:
child, employer^{-1}*, workInstitution^{-1}*

Statistical and lexical metrics for ranking
Alignments:
numberOfChildren ↔ child
staffSize ↔ employer^{-1}
staffSize ↔ workInstitution^{-1}

*Inverse predicates

Refer **Ghosh S. et al.** *Uncovering Hidden Semantics of Set Information in Knowledge Bases* CoRR abs/2003.03155 (under revision at JWS 2020) for more details.

System Description

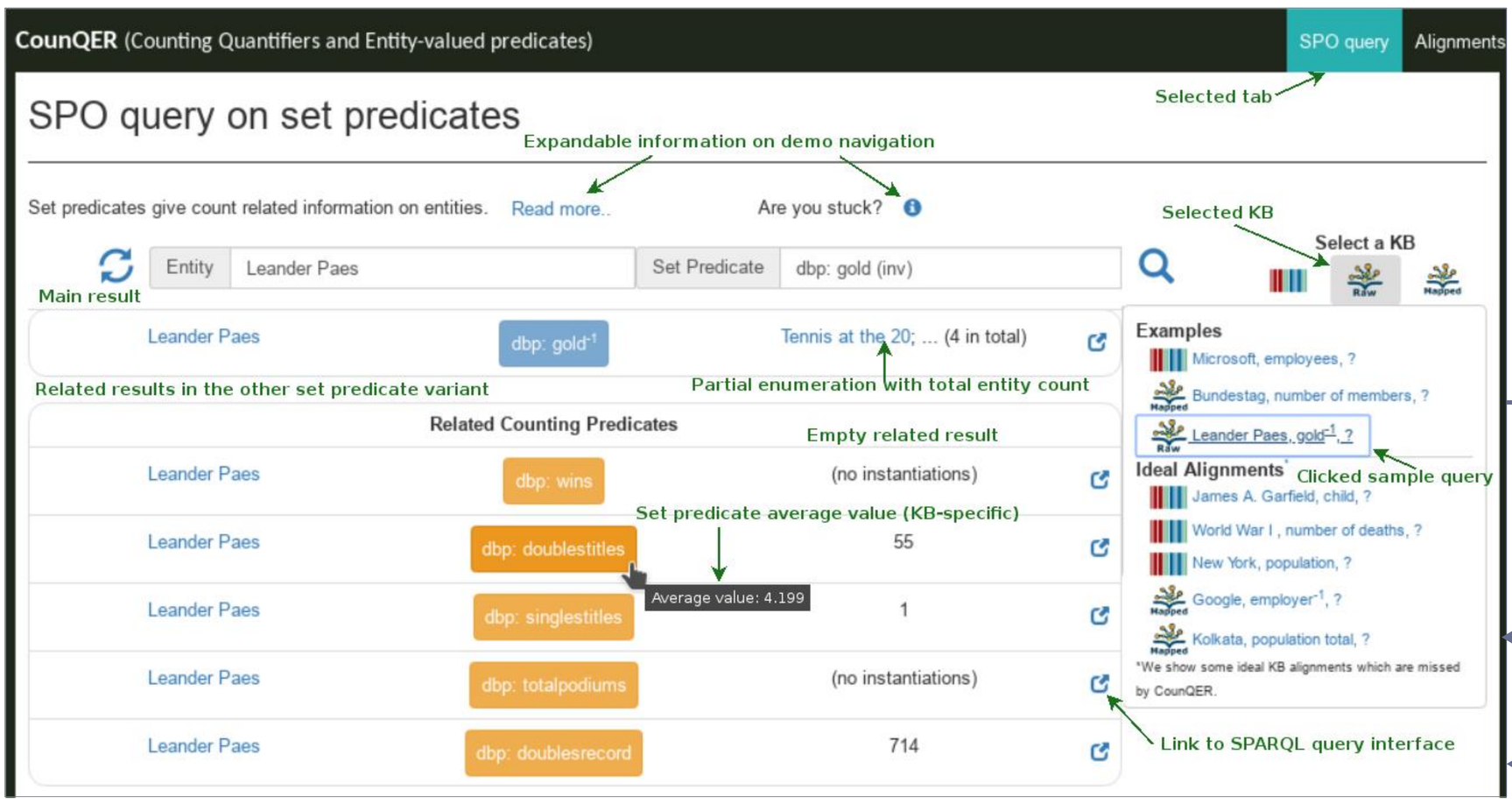
SPO Query. *Interface to query on KB-specific set predicates and entities with real-time entity suggestions.*

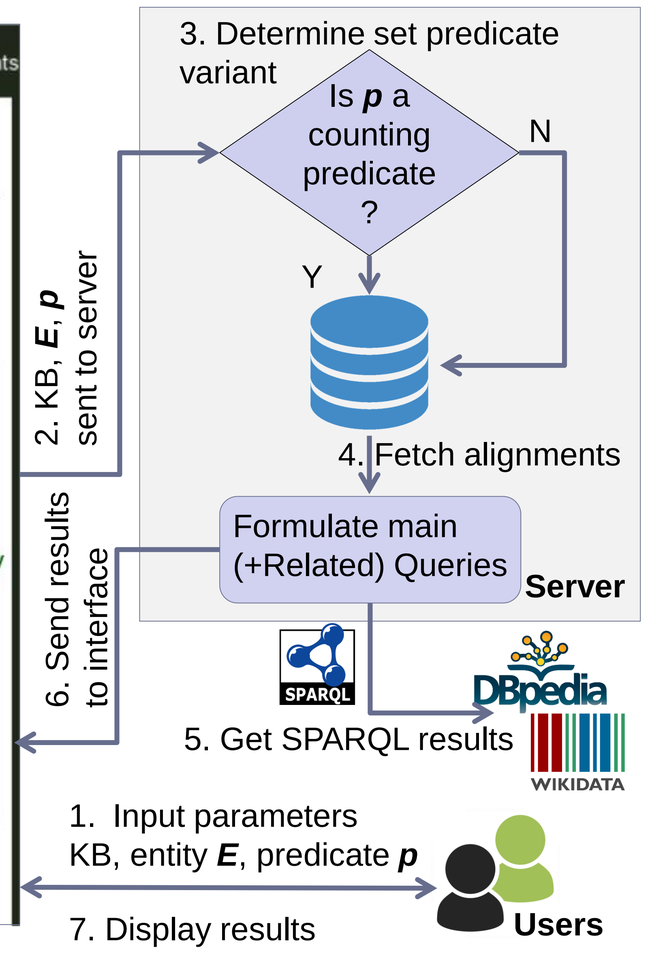
- Send user selected input parameters (KB, entity, predicate) to our server.
- Shortlist top-5 highest scoring pairs (if available).
- Fire SPARQL queries to the corresponding KB endpoint.

KB Alignments. *Interface to view all discovered alignments across Wikidata, Freebase and two DBpedia versions (raw and mapped).*

Alignments can be sorted by their scores and are searchable by their labels.

Visit us at **counqer.mpi-inf.mpg.de/spo** or scan this QR code.





System Highlights

Entity-specific set predicate suggestions, ordered based on whether they are populated and have alignments.
Empty main results are supplemented through highest-ranked and instantiated alignments.
Links to SPARQL queries can be followed to check out the actual query made by CounQER to the endpoints.
Set predicate statistics show the average integer value or the average number of entities the set predicate takes.

MAX-PLANCK-GESELLSCHAFT

max planck institut informatik