

Assignment I: Querying Semantic Web Data (Wikidata) with SPARQL

10 points

Objective

Wikidata is one of the largest knowledge graph consisting of semantic metadata about almost everything. By working on this assignment, you will be able to

- explore the structure of the Wikidata
- perform question answering on Wikidata using SPARQL query
- use Wikidata SPARQL query builder with different choices for the visualization of the result
- Converting queries into their corresponding algebraic forms.

The Task

Write 5 SPARQL queries on Wikidata considering

- You are free to choose the questions you want to get answer for
- **The queries should be non-trivial and interesting. These will be the primary quality indicators. The evaluation will be subjective.**
- The queries should cover different **complex combinations of Union Graph Pattern, Optional Graph Pattern, Usage of different filters with complex operations, Nested query, Aggregation operation with or without constraint.**
- Appropriate visualization should be used to present the result of the queries. You may achieve this by adding a comment like `#defaultView:Map` in the query.
- **The queries must be original and unique.** They should not be copied or modified from the lists of example queries in the Wikidata or any other website, or mentioned in my slides or written by your peer in the class.
- **Convert each query into their corresponding SPARQL algebra form following the TRANSFORM algorithm. Then use ARQ Jena tool to do the same. Self-report the mistakes (if any) that you made. Absolute honesty is solicited here.**

Useful Resources

- Wikidata query builder: <https://query.wikidata.org/>
- My tutorial on Wikidata, SPARQLWrapper: <https://www.youtube.com/watch?v=YAqIDLCU1Gg>
- Wikidata Query Examples:
https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/queries/examples

Submission Format

A document (pdf, doc) in the following format.

Name:

Roll No:

Queries:

- Natural language query String: Example: The list of living female Indian cricketers
- Wikidata Query: <The entire query in formatted text>
- Executable query link: <https://w.wiki/gn4>
- Your SPARQL algebra form, Apache Jena converted form and the list of mistakes you made.