

IRE Assignment 2

Problem Statement: How to create Wikipedia pages from Wikidata in your mother tongue?

Things to read before starting the assignment:

1. Wikidata
2. SPARQL for Wikidata

SPARQL is a language to formulate questions (queries) for Wikidata. You can answer queries like “*what is the most popular tonality in music?*” or “*which character was portrayed by the most actors?*” or “*what’s the distribution of blood types?*”

It can be easily integrated into python or java code. Few resources that might be helpful:

<https://query.wikidata.org/>

https://www.wikidata.org/wiki/Wikidata:SPARQL_tutorial

https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/queries/examples

https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service (Learning SPARQL section in it.)

Where to create the Wikipedia page?

Please do not create a new page in the main Wikipedia. Wikipedia has a Sandbox, which is a space to experiment with the process of editing Wikipedia. You can find information about sandbox here:

<https://en.wikipedia.org/wiki/Wikipedia:Sandbox>

Create your page in the sandbox and provide us with the sandbox link.

What needs to be done?

Create a Wikipedia page automatically about **anything** using Wikidata in your mother tongue.

The problem has two parts:

1. Query Wikidata to extract information about the item whose page you are creating.
2. Convert this information (RDF Triplets) into a Wikipedia article automatically. Your code should convert RDF Triplets into the text for the Wikipedia article. You can manually copy this generated text into the Wikipedia editor.

An example of the generated article:

<https://en.wikipedia.org/wiki/User:Himanshu.1007/sandbox>

Please note that all the editing (generating heading, table of content, bolding keywords, etc.) was done manually. Your code should only generate the content of the page, editing can be done manually.

What to submit?

A zip file (rollno.zip) containing:

1. A report that gives a brief explanation about Wikidata and SPARQL and a detailed explanation about your approach: the report should also contain a link to the Wikipedia page that you created.
2. Code

Deadline:

2 Nov 2021, 10 am (IST)

No extension, so start early.

Note: Any form of academic dishonesty will lead to 0.