

# SPARQL queries for Project Jarā'íd

## Basic documentation

Till Grallert<sup>1,2,\*</sup>

2025-12-02

<sup>1</sup> Humboldt-Universität zu Berlin

<sup>2</sup> NFDI4Memory

\* Correspondence: Till Grallert <till.grallert@hu-berlin.de>

Queries are largely modularised, which allows for some plug-and-play approach in chaining modules together. They are organised in two separate folders:

```
.  
├── queries/          # folder with SPARQL queries  
└── subqueries/       # folder for routines and sub queries
```

### file naming conventions

- file names begin with the item type
- file names use key-value pairs. keys are separated by dashes and values by underscores: `-key_value-key_value`.
  - take for example the language of periodicals, which uses the `lang` key and ISO or BCP47 language codes, i.e. `lang_ar` for Arabic.
- The type of display resulting from a SPARQL query is suffixed to the file name using the `disp` key:
  - nothing or `disp_table`: the default table view
  - `disp_bubble`: `#defaultView:BubbleChart`
  - `disp_map`: `#defaultView:Map{"hide": ["?coords"]}`
  - `disp_graph`: `#defaultView:Graph`
- reusable subqueries are marked with the prefix `subquery_`

### overview of queries

Most queries are far too long for Wikimedia's link shortener for sharing. I have therefore submitted many to the query chest service. A list of queries is provided in a separate file (`queries_list.md`).

## sample query

The following `iframe` embeds tabular view of all Arabic periodicals published before 1930 in a HTML page. Note that `iframes` are not loaded if Markdown is displayed on GitHub etc.

## General workflow

### 1. Select the basic data set

```
PREFIX medium: <http://www.wikidata.org/entity/Q1002697>      # set a publication type: periodicals (wd:Q1002697)
SELECT DISTINCT
*
# all periodicals before the date of interest
WITH {
  SELECT DISTINCT
    ?periodical ?dateOnset
  WHERE {
    VALUES ?dateOfInterest {"1910-06-01"^^xsd:dateTime}. # set a date of interest
    ?periodical wdt:P31/wdt:P279* medium: ;                  # limit to medium
    (wdt:P571 | wdt:P580) ?dateOnset.
    ?periodical wdt:P407/wdt:P279* wd:Q13955.             # limit by publication language (Arabic = Q13955)
    FILTER(?dateOnset < ?dateOfInterest).                 # published before a specific date
  }
  ORDER BY ?dateOnset
  LIMIT 4000
} as %periodicals
```

### 2. Retrieve additional information

## FILTERS

### Dates

- with data type:
  - `FILTER(?dateOnset < "1930-01-01"^^xsd:dateTime)`
- without data type but additional functions
  - `FILTER( YEAR(?dateOnset) < 1930)`
  - `FILTER( YEAR(COALESCE(?dateInception, ?dateStart)) < 1930).`

## Maps

1. Get locations: `?periodical (wdt:P291|wdt:P159|wdt:P276|wdt:P495|wdt:P131 ) ?pubPlace.`
2. Get coordinates `?pubPlace wdt:P625 ?coords`
3. Get geoshape `?pubPlace wdt:P3896 ?geoshape`
4. Options: `#defaultView:Map{"hide":["?coords"], "markercluster": "true"}`