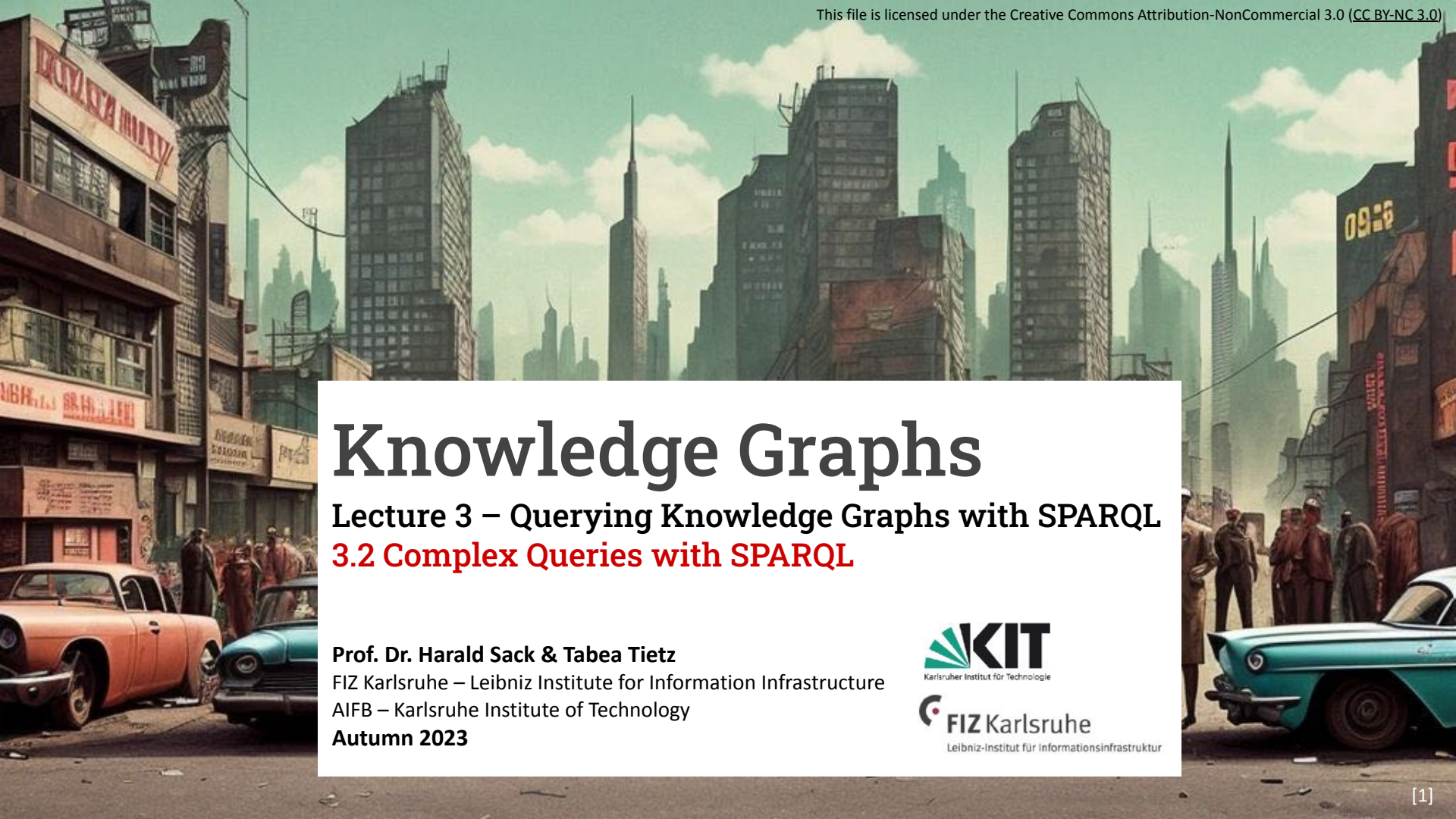


Knowledge Graphs

Lecture 3 – Querying Knowledge Graphs with SPARQL

3.2 Complex Queries with SPARQL

Prof. Dr. Harald Sack & Tabea Tietz
FIZ Karlsruhe – Leibniz Institute for Information Infrastructure
AIFB – Karlsruhe Institute of Technology
Autumn 2023

Karlsruher Institut für Technologie

FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Knowledge Graphs

Lecture 3: Querying Knowledge Graphs with SPARQL

3.1 How to Query RDF(S)

Excursion 3: DBpedia Knowledge Graph

Excursion 4: Wikidata Knowledge Graph

3.2 Complex Queries with SPARQL

3.3 More Complex SPARQL Queries

3.4 SPARQL Sub-Select and Property Paths

3.5 SPARQL is more than a Query Language

3.6 Quality Assurance with SHACL Constraints

SPARQL Filter Constraints

```
PREFIX : <http://dbpedia.org/resource/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX dct:  <http://purl.org/dc/terms/>
PREFIX dbc:  <http://dbpedia.org/resource/Category:>

SELECT ?author_name ?title
FROM <http://dbpedia.org/>
WHERE {
       ?author rdf:type dbo:Writer .
       ?author rdfs:label ?author_name
       FILTER (LANG(?author_name)="en").
       ?work dbo:author ?author .
             ?work rdfs:label ?title .
       FILTER (LANG(?title)="en")
        ?work dct:subject dbc:Dystopian_novel .
} LIMIT 100
```

Example:
Search for **authors** and their **books**, filter results for **English labels** and **Dystopian novels** and **limit** the results to **the first 100.**

DBpedia

query SPARQL endpoint

SPARQL Filter Constraints

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX wikibase: <http://wikiba.se/ontology#>
PREFIX bd: <http://www.bigdata.com/rdf#>

SELECT ?authorLabel ?bookLabel ?date
WHERE {
      ?book wdt:P31 wd:Q7725634 .   # ?book :instanceOf :LiteraryWork
      ?book wdt:P50 ?author .       # ?book :author ?author
      ?book wdt:P136 wd:Q15062348 . # ?book :genre :DystopianFiction
      ?book wdt:P577 ?date .        # ?book :publicationDate -?date
      SERVICE wikibase:label
      { bd:serviceParam wikibase:language "en" }
} LIMIT 100
```

Example:
Search for **authors** and their **books** including **publication date**, filter results for **English labels** and **Dystopian novels** and **limit** the results to **the first 100.**

WIKIDATA

query SPARQL endpoint

SPARQL Filter Constraints

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX wikibase: <http://wikiba.se/ontology#>
PREFIX bd: <http://www.bigdata.com/rdf#>

SELECT ?authorLabel ?bookLabel ?date
WHERE {
      ?book wdt:P31 wd:Q7725634 .   # ?book :instanceOf :LiteraryWork
      ?book wdt:P50 ?author .       # ?book :author ?author
      ?book wdt:P136 wd:Q15062348 . # ?book :genre :DystopianFiction
      ?book wdt:P577 ?date .        # ?book :publicationDate -?date
      SERVICE wikibase:label
      { bd:serviceParam wikibase:language "en" }
} Limit 100
```

wikidata specific label service

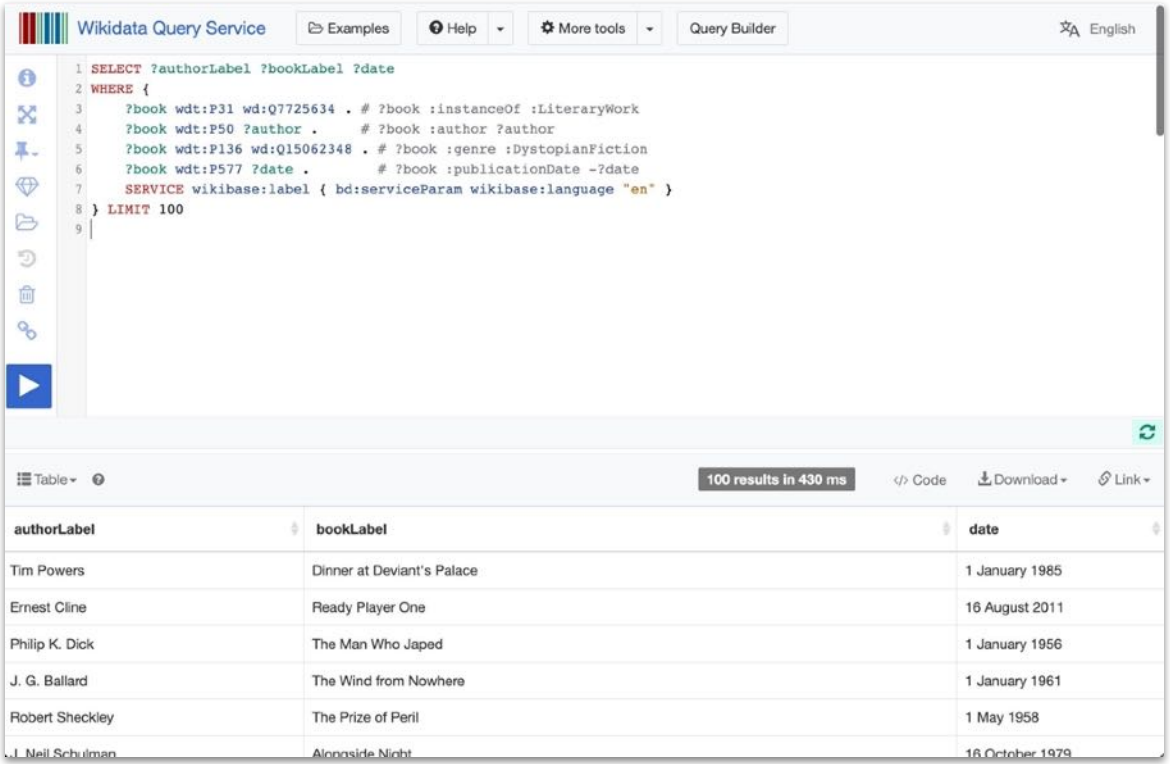
Example:
Search for **authors** and their **books** including **publication date**, filter results for **English labels** and **Dystopian novels** and **limit** the results to **the first 100.**

WIKIDATA

query SPARQL endpoint

Example:

Search for **authors** and their **books** including **publication date**, filter results for **English labels** and **Dystopian novels** and **limit** the results to **the first 100.**



query SPARQL endpoint

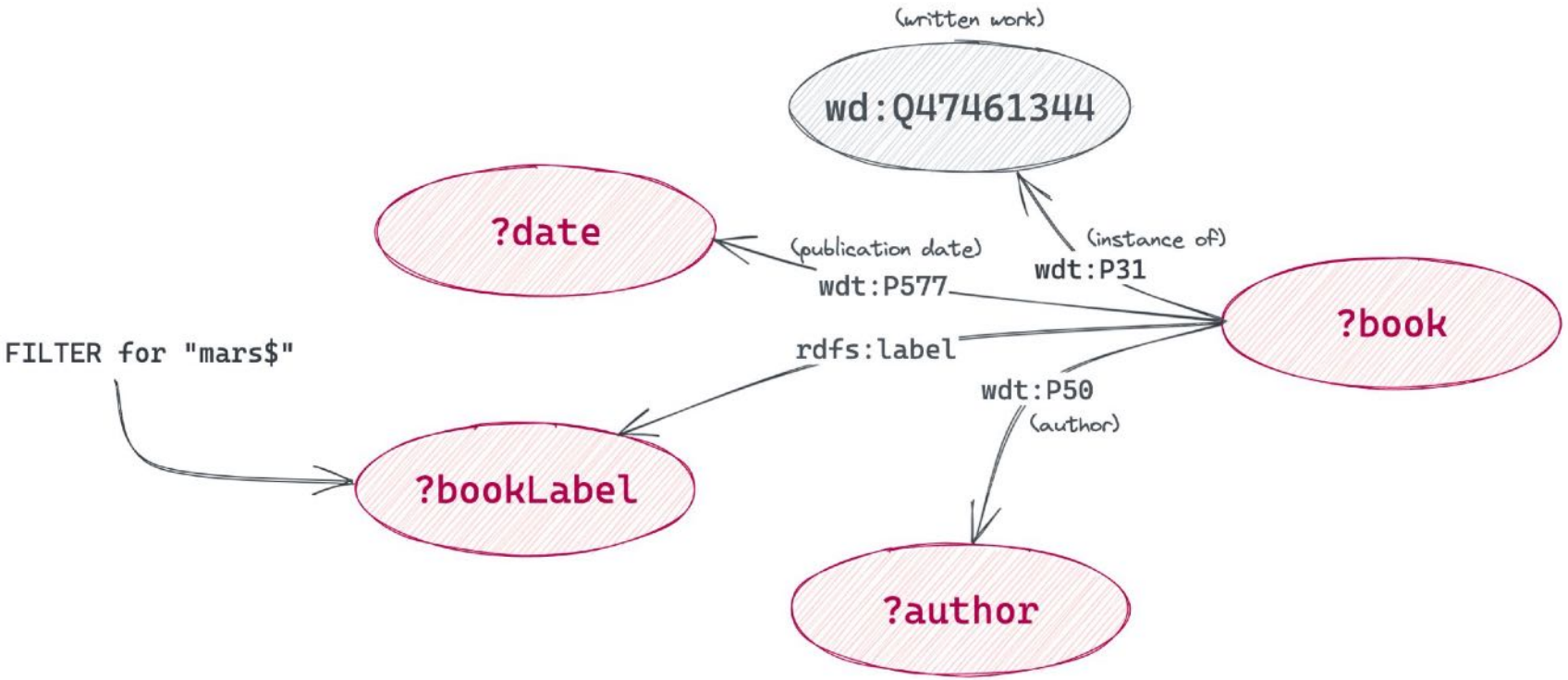
More SPARQL Operators

- Logical connectives `&&` (**AND**) and `||` (**OR**) for xsd:boolean
- Comparison operators `=`, `!=`, `<`, `>`, `<=`, and `>=` for numeric data types, xsd:dateTime, xsd:string, and xsd:boolean
- Comparison operators `=` and `!=` for other data types
- Arithmetic operators `+`, `-`, `*`, and `/` for numeric data types

- And in addition:
  - `REGEX(String,Pattern)` or `REGEX(String,Pattern,Flags)`
  - `sameTERM(A,B)`
  - `langMATCHES(A,B)`

SPARQL Filter Constraints

Search for book titles that **end with the word "mars"** sorted by publication date?



SPARQL Filter Constraints

Search for book titles that **end with the word "mars"** sorted by publication date?

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT ?authorLabel ?bookLabel ?date
WHERE {
      ?book wdt:P31 wd:Q47461344 .
      ?book wdt:P50 ?author .
      ?book wdt:P577 ?date .
     ?book rdfs:label ?bookLabel
      FILTER (LANG(?bookLabel)="en")
     FILTER REGEX (?bookLabel,"mars$","i") .
     ?author rdfs:label ?authorLabel
      FILTER (LANG(?authorLabel)="en") .
} ORDER BY ?date
```

With **FILTER REGEX**, regular expressions can be filtered.

query SPARQL endpoint

SPARQL Filter Constraints

Search for book titles that **end with the word "mars"** sorted by publication date?

Wikidata Query Service | Examples | Help | More tools | Query Builder | English

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT ?authorLabel ?bookLabel ?date
WHERE {
    ?book wdt:P31 wd:Q47461344 .
    ?book wdt:P50 ?author .
    ?book wdt:P577 ?date .
    ?book rdfs:label ?bookLabel
    FILTER (LANG(?bookLabel)="en")
    FILTER REGEX (?bookLabel,"mars$","i") .
    ?author rdfs:label ?authorLabel .
FILTER (LANG(?authorLabel)="en")
} ORDER BY ?date
```

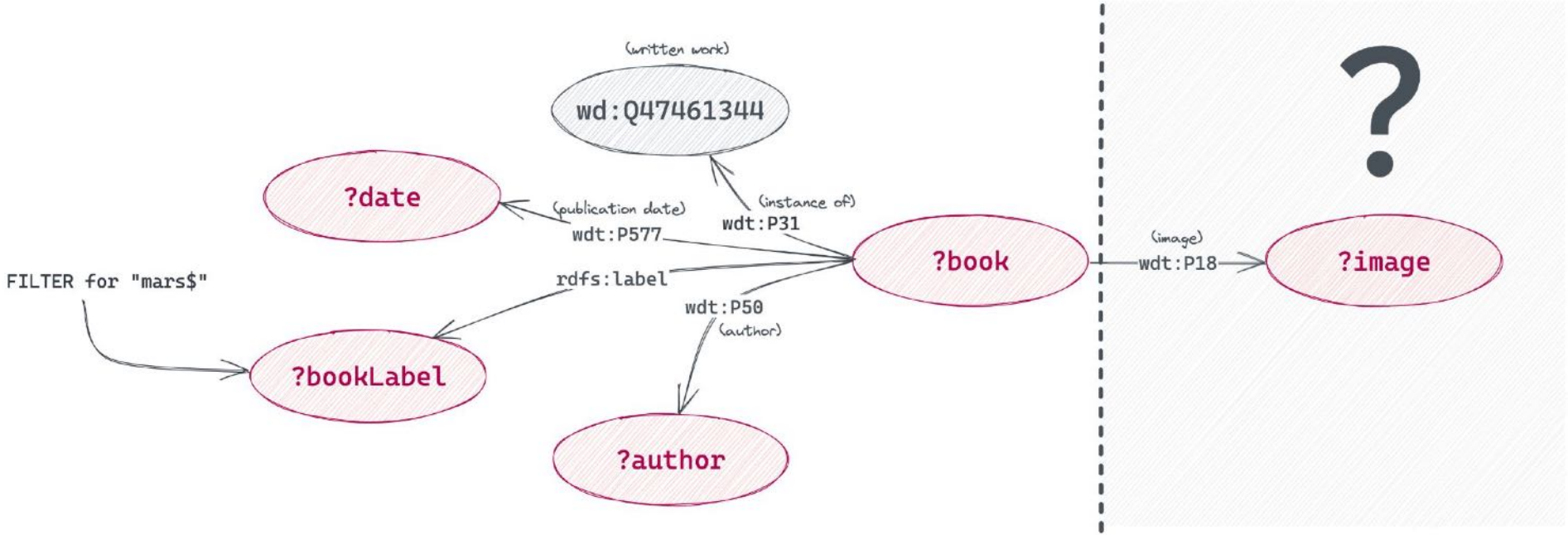
Table | 18 results in 5470 ms | Code | Download | Link

authorLabel	bookLabel	date
Alfred de Vigny	Cinq-Mars	25 April 1826
Edgar Rice Burroughs	Thuvia, Maid of Mars	1 January 1920
Edgar Rice Burroughs	A Fighting Man of Mars	1 January 1930
Edgar Rice Burroughs	Synthetic Men of Mars	1 January 1940
Garrett P. Serviss	Edison's Conquest of Mars	1 January 1947
Robert A. Heinlein	Podkayne of Mars	1 January 1963
Edgar Rice Burroughs	John Carter of Mars	1 January 1964

query SPARQL endpoint

SPARQL Optional Constraints

Search for book titles that **end with the word "mars"**, and **optionally have an image**?



SPARQL Optional Constraints

Search for book titles that **end with the word "mars"**, and **optionally have an image**?

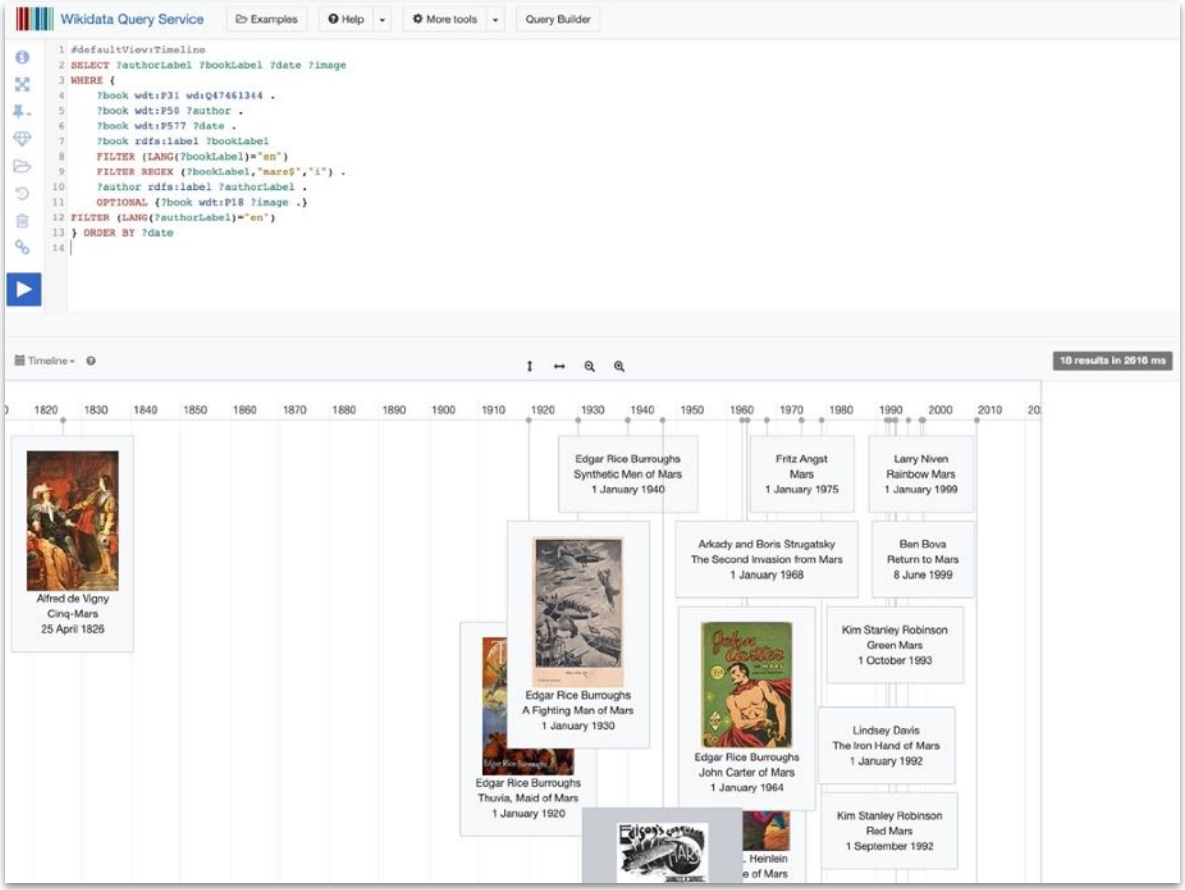
```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT ?authorLabel ?bookLabel ?date ?image
WHERE {
      ?book wdt:P31 wd:Q47461344 .
      ?book wdt:P50 ?author .
      ?book wdt:P577 ?date .
     ?book rdfs:label ?bookLabel
      FILTER (LANG(?bookLabel)="en")
     FILTER regex (?bookLabel,"mars$","i") .
     ?author rdfs:label ?authorLabel
      FILTER (LANG(?authorLabel)="en") .
      OPTIONAL {?book wdt:P18 ?image}
} ORDER BY ?date
```

optional constraint

Optional selection of graph pattern via **OPTIONAL**.

query SPARQL endpoint

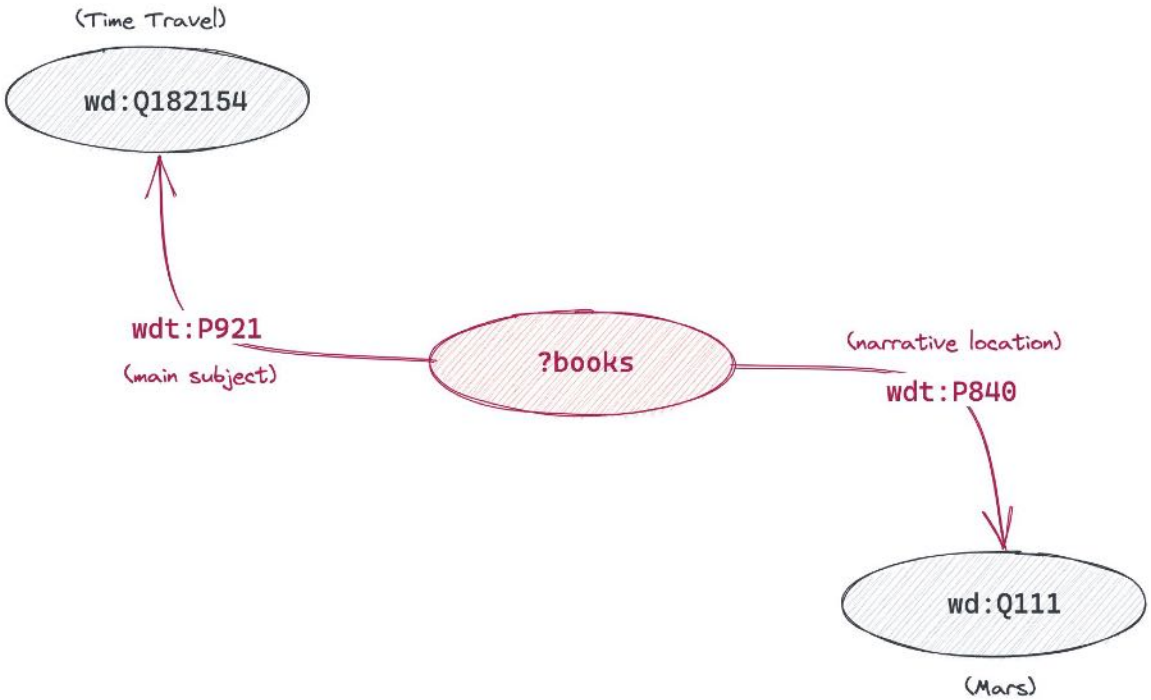
Search for book titles that **end with the word "mars"**, and **optionally have an image**?



[query SPARQL endpoint]

SPARQL Alternative Results via UNION

Example: which books are dealing with time travel or have Mars as their narrative location?



SPARQL Alternative Results via UNION

Example: which books are dealing with time travel or have Mars as their narrative location?

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX wikibase: <http://wikiba.se/ontology#>
PREFIX bd: <http://www.bigdata.com/rdf#>

SELECT ?authorLabel ?bookLabel ?date
WHERE {
      ?book wdt:P31 wd:Q47461344 ;        # ?book :instanceof :LiteraryWork
            wdt:P50 ?author ;             # :author ?author
            wdt:P577 ?date .              # :publicationDate ?date
     { ?book wdt:P921 wd:Q182154 . }      # ?book :mainSubject :TimeTravel
    UNION
     { ?book wdt:P840 wd:Q111 .}          # ?book :narrativeLocation :Mars
  SERVICE wikibase:label
  { bd:serviceParam wikibase:language "en,fr,it"} # labeling service
} ORDER by ?date
```

logical disjunction

The keyword **UNION** allows for alternatives (logical disjunction).

query SPARQL endpoint

SPARQL Alternative Results via UNION

Example: which books are dealing with time travel or have Mars as their narrative location?

Wikidata Query Service | Examples | Help | More tools | Query Builder

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX wikibase: <http://wikiba.se/ontology#>
PREFIX bd: <http://www.bigdata.com/rdf#>

SELECT ?authorLabel ?bookLabel ?date
WHERE {
    ?book wdt:P31 wd:Q47461344 ;  # ?book :instanceof :LiteraryWork
          wdt:P50 ?author ;       # :author ?author
          wdt:P577 ?date .        # :publicationDate ?date
    { ?book wdt:P921 wd:Q182154 . }  # ?book :mainSubject :TimeTravel
    UNION
    { ?book wdt:P840 wd:Q111 .}      # ?book :narrativeLocation :Mars
  SERVICE wikibase:label { bd:serviceParam wikibase:language "en, fr, it" } # labeling service
} ORDER by ?date
```

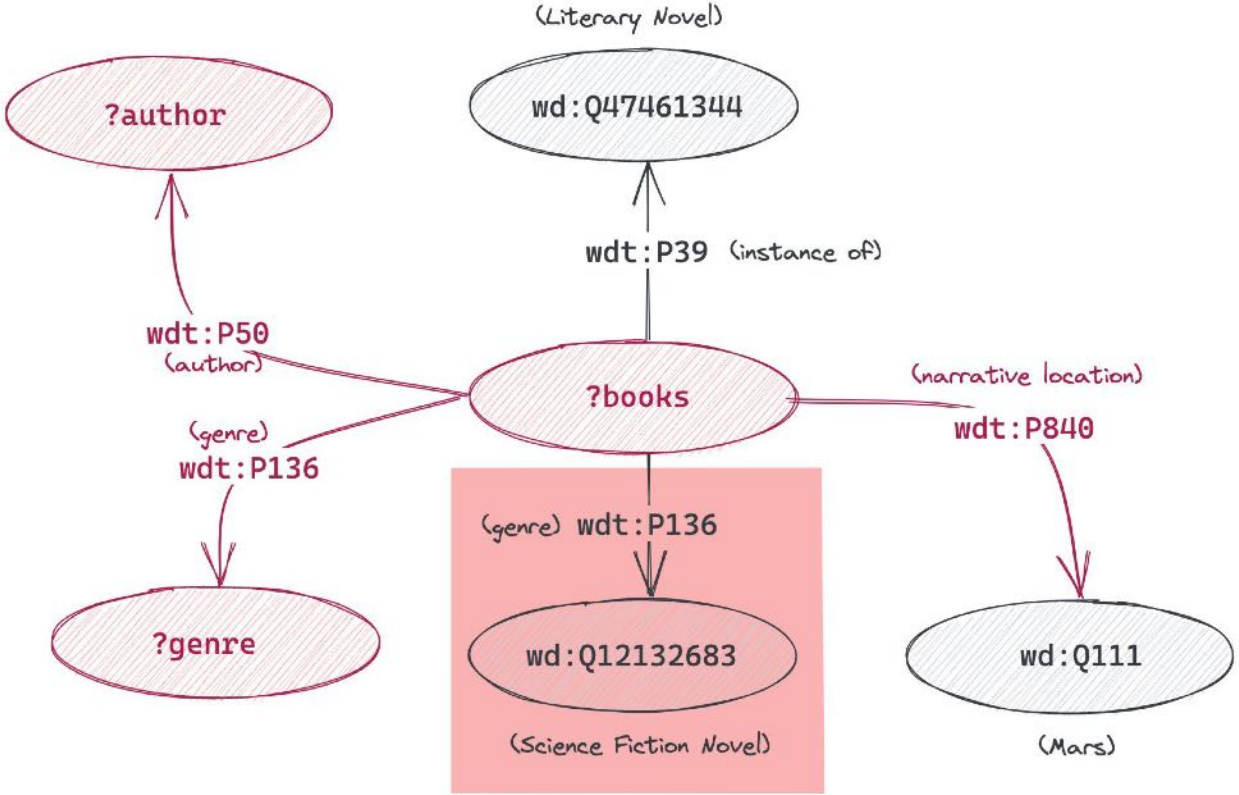
Table | 30 results in 349 ms | Code

authorLabel	bookLabel	date
Emilio Salgari	Le meraviglie del Duemila	1 January 1907
Edgar Rice Burroughs	A Fighting Man of Mars	1 January 1930
Wilson Tucker	The Lincoln Hunters	1 January 1958
Robert A. Heinlein	Podkayne of Mars	1 January 1963
Robert A. Heinlein	Farnham's Freehold	1 January 1964
Robert Silverberg	The Time Hoppers	1 January 1967
Jack Finney	The Woodrow Wilson Dime	1 January 1968
Robert Silverberg	The Masks of Time	1 May 1968

query SPARQL endpoint

SPARQL Negation

Example: which books have Mars as their narrative location and are not Science Fiction novels?



SPARQL Negation

Example: which books have Mars as their narrative location and are not Science Fiction novels?

```
SELECT ?authorLabel ?bookLabel ?genreLabel
WHERE {
     ?book wdt:P31 wd:Q47461344 ; # ?book :instanceof :LiteraryWork
           wdt:P50 ?author ;      # :author ?author
           wdt:P840 wd:Q111 ;     # :narrativeLocation :Mars
           wdt:P136 ?genre .      # :genre ?Genre
    FILTER NOT EXISTS {?book wdt:P136 wd:Q12132683 }.
                      # ?book :genre :ScienceFictionNovel
  SERVICE wikibase:label { bd:serviceParam wikibase:language "en" }
}
```

filter query result for existence

SPARQL 1.1 offers several variants for negation:

- **FILTER NOT EXISTS**
- **MINUS**
- **!BOUND()**

query SPARQL endpoint

SPARQL Negation

Example: which books have Mars as their narrative location and are not Science Fiction novels?

Wikidata Query Service | Examples | Help | More tools | Query Builder

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX wikibase: <http://wikiba.se/ontology#>
PREFIX bd: <http://www.bigdata.com/rdf#>

SELECT ?authorLabel ?bookLabel ?genreLabel
WHERE {
    ?book wdt:P31 wd:Q47461344 ;  # ?book :instanceof :LiteraryWork
          wdt:P50 ?author ;       # :author ?author
          wdt:P840 wd:Q111 ;      # :narrativeLocation :Mars
          wdt:P136 ?genre .       # :genre ?Genre
   FILTER NOT EXISTS {?book wdt:P136 wd:Q12132683 }. # ?book :genre :ScienceFictionNovel
  SERVICE wikibase:label { bd:serviceParam wikibase:language "en" } # labeling service
}
```

Table | 4 results in 286 ms

authorLabel	bookLabel	genreLabel
Robert A. Heinlein	Podkayne of Mars	science fiction
Diane Duane	A Wizard of Mars	speculative/fantastic fiction novel
Larry Niven	Rainbow Mars	horror literature
Larry Niven	Rainbow Mars	crossover fiction

query SPARQL endpoint

Next Lecture...

[3]

Bibliographic References:

- Steve Harris, Andy Seaborne (2013), *SPARQL 1.1 Query Language*, W3C Recommendation 21 March 2013
- Aidan Hogan (2020), *The Web of Data*, Springer.
  - Chap. 6.2.2 Unions of Graph Patterns, pp. 333–336.
  - Chap, 6.2.3 Optional Graph Patterns, pp. 336–338.
  - Chap. 6.2.4 Filtering and Binding Patterns, pp. 338–346.
  - Chap. 6.2.5 Negation of Graph Patterns, pp. 346–351.

Picture References:

[1] "A dystopian city street scene clearly exhibiting the consequences of both unchecked population growth on society and the hoarding of resources by a wealthy minority in the style of a 1960s pulp cover." , created via ArtBot, Deliberate, 2023, [CC-BY-4.0], https://tinybots.net/artbot

[2] DBpedia logo, wiki.dbpedia.org, DBpedia Team [Public Domain], https://commons.wikimedia.org/wiki/File:DBpediaLogo.svg

[3] The Linked Open Data Cloud, lod-cloud.net, [CC-BY], https://lod-cloud.net/clouds/lod-cloud.svg

[4] "In this 1960s pulp cover picture, in the waning days of a future Galactic Empire, the mathematician Hari Seldon spends his life developing a theory of psychohistory, a new and effective mathematics of sociology. Using statistical laws of mass action, it can predict the future of large populations.", created via ArtBot, Deliberate, 2023, [CC-BY-4.0], https://tinybots.net/artbot