

Tutorial: SPARQL

Wie kann ich SPARQL praktisch anwenden? In diesem Tutorial werden Sie einfache bis komplexe SPARQL Abfragen schreiben und dabei verschiedene Datensätze verlinken. Für das Tutorial sind keine Programmierkenntnisse erforderlich.

Das Tutorial führt Sie durch folgende Schritte:

- Komplexe SPARQL Abfragen bauen in SPARKLIS
- Einen lokalen SPARQL Endpoint mit Jena-Fuseki einrichten
- SPARQL Abfragen schreiben und Daten verlinken

Glossar und verwendete Akronyme

HTTP: Hypertext Transfer Protocol

OWL: Web Ontology Language

RDF: Resource Description Framework

SPARQL: SPARQL Protocol And RDF Query Language

URI: Uniform Resource Identifier

W3C: World Wide Web Consortium

SPARQL Syntax

```
PREFIX ex: <http://example.org/ >
SELECT DISTINCT ?s ?p ?o WHERE {
  ?s ?p ?o
}
LIMIT 10
```

SPARKLIS

Öffnen Sie den folgenden Link: <http://www.irisa.fr/LIS/ferre/sparklis/examples.html>

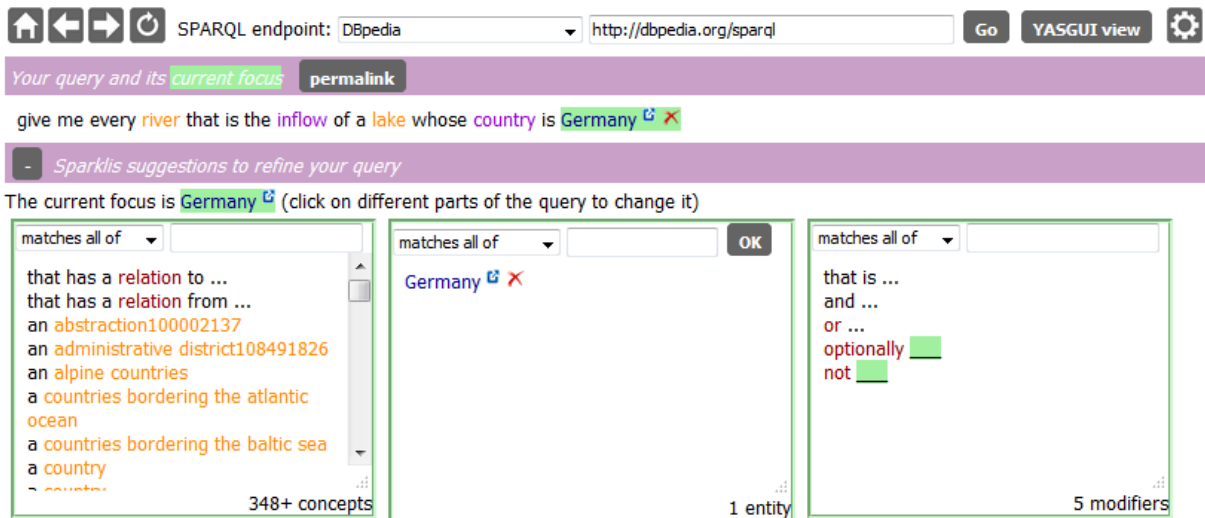
Wählen Sie ein Beispiel aus:

8. *Which rivers flow into a German lake?* -> [Open in Sparklis](#)

Wechseln sie zur *YASGUI view* (in der Menüleiste neben *Go*) um die entsprechende SPARQL Abfrage zu sehen.

Versuchen Sie die Abfrage nach eigenem Wunsch abzuändern.

Beispiel: Ersetzen Sie zum Beispiel den Ländernamen mit „dbr:Switzerland“. Probieren Sie aus.



SPARQL endpoint: DBpedia

Your query and its **current focus**

give me every river that is the inflow of a lake whose country is Germany

Sparklis suggestions to refine your query

The current focus is **Germany** (click on different parts of the query to change it)

matches all of that has a relation to ... that has a relation from ... an abstraction100002137 an administrative district108491826 an alpine countries a countries bordering the atlantic ocean a countries bordering the baltic sea a country ~ countries 348+ concepts	matches all of <input type="text" value="Germany"/> <input type="button" value="OK"/> Germany	matches all of that is ... and ... or ... optionally not
--	--	---

SPARQL Abfrage:

```

PREFIX dbr: <http://dbpedia.org/resource/>
PREFIX dbo: <http://dbpedia.org/ontology/>
SELECT DISTINCT ?River_1 ?Lake_2
WHERE { ?River_1 a dbo:River .
        ?Lake_2 a dbo:Lake .
        ?Lake_2 dbo:inflow ?River_1 .
        ?Lake_2 dbo:country dbr:Switzerland . }
LIMIT 200
  
```

Testen Sie Ihre Abfrage im dem SPARQL Endpoint von DBpedia: <http://dbpedia.org/sparql>

Jena-Fuseki SPARQL Endpoint installieren

1. Letzte Version von Fuseki herunterladen:

<http://www.apache.org/dist/jena/binaries/> -> apache-jena-fuseki-2.3.0.zip

2. Datei entpacken

3. Datei auf dem Desktop ablegen

4. Eingabe in Kommandozeile

- Windows (Programme > Zubehör > Eingabeaufforderung):

D:	Press enter
cd apache-jena-fuseki-2.3.0	Press enter
fuseki-server --update --mem /ds	Press enter

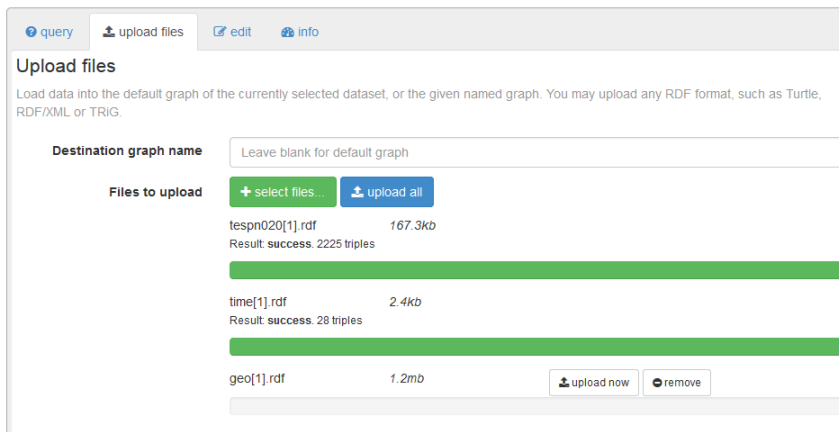
- Mac (Programme > Dienstprogramme > Terminal):

cd Desktop/apache-jena-fuseki-2.3.0	Press enter
./fuseki-server --update --mem /ds	Press enter

5. Im Web-Browser folgende Adresse aufrufen: <http://localhost:3030> (<http://localhost:3030>)

Sie haben jetzt Ihren eigenen lokalen SPARQL Endpoint und RDF Triple Store eingerichtet.

Gehen Sie zum Menüpunkt: add data -> upload files



Laden Sie die folgenden Daten von EUROSTAT hoch (+select files..., Link-Adresse einsetzen, öffnen):

- Datensatz mit Kennzahlen (Observations): <http://eurostat.linked-statistics.org/data/> -> Indikator auswählen -> z.B. [tespn020.rdf](http://eurostat.linked-statistics.org/data/tespn020.rdf) („Median relative income of elderly people“) -> mit der rechten Maustaste ‚Link-Adresse kopieren...‘
- Geo Dimension: <http://eurostat.linked-statistics.org/dic/time.rdf>
- Time Dimension: <http://eurostat.linked-statistics.org/dic/geo.rdf>

-> Upload now

Wechseln Sie zum Menüpunkt ‚query‘ und schreiben Sie Ihre erste SPARQL Abfrage.

```

SELECT DISTINCT ?s ?p ?o WHERE {
  ?s ?p ?o
}
LIMIT 100
  
```

Sie sehen nun die ersten 100 Triples.

Mit der nächsten Abfrage werden die im Datensatz „tespn020“ enthaltenen Kennzahlen zu jedem Land abgefragt (hier „Median relative income of elderly people“). Dieselbe Abfrage funktioniert auch mit anderen Indikatoren von EUROSTAT Daten.

```

# This is a sample SPARQL query that returns the first 100 countries
# included in the EUROSTAT data for indicator "tespn020" and year "2010"

PREFIX qb: <http://purl.org/linked-data/cube#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX dad-prop: <http://semantic.digital-agenda-data.eu/def/property/>
PREFIX sdmx-measure: <http://purl.org/linked-data/sdmx/2009/measure#>
PREFIX sdmx-dimension: <http://purl.org/linked-data/sdmx/2009/dimension#>
PREFIX property: <http://eurostat.linked-statistics.org/property#>
  
```

```
SELECT DISTINCT ?dsd ?country ?countryname ?obsEUval
WHERE {
  ?obsEU qb:dataSet ?dsd .
  ?obsEU sdmx-measure:obsValue ?obsEUval .
  ?obsEU property:geo ?country .
  ?country skos:prefLabel ?countryname .
  ?obsEU sdmx-dimension:timePeriod "2010-01-01" ^^<http://www.w3.org/2001/XMLSchema#date> .
}
ORDER BY ?country
LIMIT 100
```

Wir können nun mit Hilfe der ‚SERVICE‘ Funktion einen externen SPARQL Endpoint und so Informationen von anderen Datensets abfragen (z.B. Digital Agenda Data und den Indikator i_iu3: „Individuals who have used internet in the last 3 months“, Breakdown: All individuals aged 16-74).

```
# This is a sample SPARQL (SERVICE) query that returns the first 100 countries
# included in the Digital Agenda Data for indicator "i_iu3" and year "2010"

PREFIX qb: <http://purl.org/linked-data/cube#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX dad-prop: <http://semantic.digital-agenda-data.eu/def/property/>
PREFIX sdmx-measure: <http://purl.org/linked-data/sdmx/2009/measure#>

SELECT DISTINCT ?country ?countryname ?obsDAvalue
WHERE {
  SERVICE <http://digital-agenda-data.eu/data/sparql> {
    SELECT DISTINCT ?country ?countryname ?obsDAvalue
    WHERE {
      ?obsDA a qb:Observation .
      ?obsDA dad-prop:indicator <http://semantic.digital-agenda-data.eu/codelist/indicator/i_iu3> .
      ?obsDA dad-prop:time-period <http://reference.data.gov.uk/id/gregorian-year/2010> .
      ?obsDA dad-prop:ref-area ?country .
      ?country skos:prefLabel ?countryname .
      ?obsDA sdmx-measure:obsValue ?obsDAvalue .
      ?obsDA dad-prop:breakdown <http://semantic.digital-agenda-data.eu/codelist/breakdown/ind_total> .
    }
  }
}
ORDER BY ?country
LIMIT 100
```

Dank der verlinkten Geo-Locations können wir die Datensätze auch mit SPARQL verlinken.

“Linking the Observations”

```
# This is a sample linking SPARQL query that returns the first 100 countries
# included in the EUROSTAT data for indicator "tespn020" and year "2010"
# and the Digital Agenda Data for indicator "i_iu3" and year "2010"
```

```
PREFIX qb: <http://purl.org/linked-data/cube#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX dad-prop: <http://semantic.digital-agenda-data.eu/def/property/>
PREFIX sdmx-measure: <http://purl.org/linked-data/sdmx/2009/measure#>
PREFIX sdmx-dimension: <http://purl.org/linked-data/sdmx/2009/dimension#>
PREFIX property: <http://eurostat.linked-statistics.org/property#>

SELECT DISTINCT ?country ?countryname ?obsEUvalue ?obsDAvalue
WHERE {
  ?obsEU qb:dataSet ?dsd .
  ?obsEU sdmx-measure:obsValue ?obsEUvalue .
  ?obsEU property:geo ?country .
  ?obsEU sdmx-dimension:timePeriod "2010-01-01" ^^<http://www.w3.org/2001/XMLSchema#date> .
  SERVICE <http://digital-agenda-data.eu/data/sparql> {
    SELECT DISTINCT ?country ?countryname ?obsDAvalue
    WHERE {
      ?obsDA a qb:Observation .
      ?obsDA dad-prop:indicator <http://semantic.digital-agenda-data.eu/codelist/indicator/i_iu3> .
      ?obsDA dad-prop:time-period <http://reference.data.gov.uk/id/gregorian-year/2010> .
      ?obsDA dad-prop:ref-area ?country .
      ?country skos:prefLabel ?countryname .
      ?obsDA sdmx-measure:obsValue ?obsDAvalue .
      ?obsDA dad-prop:breakdown <http://semantic.digital-agenda-data.eu/codelist/breakdown/ind_total> .
    }
  }
}
ORDER BY ?country
LIMIT 100
```

Sie können auch einen anderen Digital-Agenda-Scoreboard-Key-Indicator auswählen und die SPARQL Abfrage entsprechend anpassen (Breakdown beachten und ebenfalls anpassen): <http://digital-agenda-data.eu/data/observations>

Dataset metadata: http://digital-agenda-data.eu/datasets/digital_agenda_scoreboard_key_indicators

Links & Tools:

- SPARQL in 11 minutes - Video: <https://www.youtube.com/watch?v=FvGndkpa4K0>
- Learning SPARQL, Bob DuCharme: [Homepage](#), [Safari Books Online](#)
- SPARQL Tutorial, Jerven Bolleman: http://people.isb-sib.ch/Jerven.Tjalling.Bolleman/Learning%20Sparql%202012_12.pdf
- SPARQL Playground: <http://sparql-playground.nextprot.org/>
- LOD Laundromat: <http://lodlaundromat.org/>

Quellen:

- <http://wiki.lod2.eu/display/LOD2DOC/Digital+Agenda+Scoreboard+linking+example>
- https://jena.apache.org/documentation/serving_data/