Домашно бр. 2 SPARQL

203235 - Димитрија Тимески

10-Nov-2023

А) Домашна задача

1. IIITO e SPARQL?

Прашален јазик (Query Language) за структурирани податоци на вебот, т.е. RDF податоци, за разлика од SQL кој е наменет за релациони податоци.

Со други зборови SPARQL е прашален јазик за Семантичкиот Web и Linked Data.

2. Кои типови на SPARQL прашања постојат?

SELECT – Враќа податоци во табеларен формат.

DESCRIBE – Враќа податоци во RDF граф формат.

ASK – Враќа true или false, доколку постои барем една тројка.

CONSTRUCT – Создава нов RDF граф.

UPDATE — Овозможува користење на сите CRUD операции само доколку има доволно пермисии.

3. Во каква логичка релација се наоѓаат меѓусебно triple pattern-ите кои се наведени во стандарден WHERE дел од SPARQL прашање?

Тие се наоѓаат во логичка релација **И** т.е. **AND**. Задолжително сите мора да бидат исполнети за да се вчита одреден ресурс.

4. Како можеме да означиме во едно SPARQL прашање дека бараниот triple pattern не е задолжителен при пребарувањето?

За да направиме SPARQL прашање кое што има не задолжителен triple pattern (тројка), го кориситиме OPTIONAL блокот:

Пример:

```
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
prefix foaf: <http://xmlns.com/foaf/0.1/>

select ?name ?url
where {
    ?person foaf:name ?name .
    OPTIONAL { ?person rdfs:seeAlso ?url }
}
```

5. Дали еден SPARQL endpoint може да се користи и од страна на апликации? Објаснете.

SPARQL endpoint-от, не само што може да се користи од страна на апликации, туку тој е и наменет за тоа.

SPARQL endpoint-от може да се праќа во GET или POST барања, а со таа функционалност тој станува моќен Restful сервис.

Б) Вовед

DBpedia (http://www.dbpedia.org) е проект кој ги извлекува структурираните информации од Wikipedia, семантички ги анотира (со RDF) и ги објавува како јавно достапни информации. DBpedia дозволува овие податоци да се поврзуваат со други RDF податоци низ WWW, да се користат за сопствени бази на знаење, а дозволува и поставување на SPARQL прашања врз податоците со кои располага.

За да го овозможи последново, DBpedia нуди пристапна точка за пишување SPARQL прашања преку веб, преку нивниот SPARQL Endpoint, поставен на: http://dbpedia.org/sparql.

<u>Напомена</u>: Доколку некогаш имате проблеми со SPARQL Endpoint-от, можете да го користите SPARQL Explorer-от, достапен на: http://dbpedia.org/snorql. Доколку DBpedia е недостапна во моментот, нејзините податоци можете да ги пребарувате и преку http://lod.openlinksw.com/sparql.

В) Практична задача

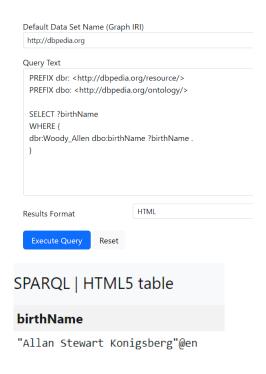
І. Работа со податоци за личност, режисер

- 1. DBреdia своите ресурси ги означува со URI-а кои се изведени од оригиналните адреси на англиските страни на Wikipedia. Тие се добиваат на следниот начин:
 - a. Се одбира Wikipedia страна (на пр: http://en.wikipedia.org/wiki/Woody_Allen)
 - b. Префиксот http://en.wikipedia.org/wiki/ се заменува со префиксот http://dbpedia.org/resource/
 - с. На тој начин, се добива DBpedia URI-то за ресурсот: http://dbpedia.org/resource/Woody_Allen
 - d. Доколку го отворите овој ресурс преку интернет прелистувач, ќе ви се отвори страната http://dbpedia.org/page/Woody_Allen, на која се излистани сите RDF тројки кои DBpedia ги извлекла од Wikipedia за дадениот ресурс.
 - е. Првата половина RDF тројки се оние во кои ресурсот се среќава како Subject. Втората половина RDF тројки, во кои релациите се означени како 'is *relation* of', се оние во кои ресурсот се среќава како Object.
- 2. Разгледајте ги тројките (релациите и вредностите) кои ги има на DBpedia за ресурсот http://dbpedia.org/resource/Woody_Allen. Секоја од релациите е означена линк кој носи кон уникатно URI кое ја идентификува релацијата и на кое е поставена нејзината дефиниција.

Преку SPARQL Endpoint-от (поставен на адресата http://dbpedia.org/sparql), напишете SPARQL прашање со кое ќе го добиете вистинското име на Вуди Ален (dbo:birthName).

<u>Напомена</u>: На линкот Namespace Prefixes (поставен на SPARQL Endpoint страната) можете да ги погледнете префиксите кои може да ги користите кај ресурсите од различни извори, без да ги наведувате експлицитно во самото SPARQL прашање. На пример, ресурсот http://dbpedia.org/resource/Woody_Allen може да се наведе како **dbr:Woody_Allen**, а

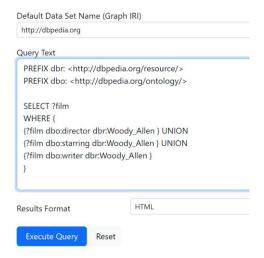
ресурсот http://dbpedia.org/ontology/birthName како dbo:birthName во SPARQL прашањата, без да ги вклучите овие два префикси експлицитно. Доколку префиксот кој сакате да го користите не е прикажан во Namespace Prefixes, дефинирајте го експлицитно.

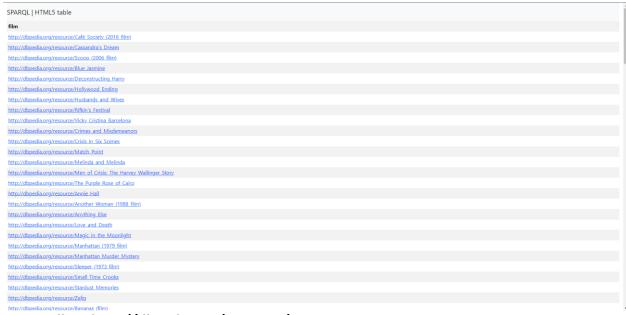


PREFIX dbr: http://dbpedia.org/ontology/>

SELECT ?birthName WHERE { dbr:Woody_Allen dbo:birthName ?birthName . }

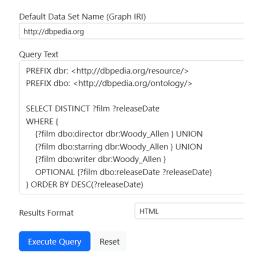
3. Напишете SPARQL прашање со кое ќе ги добиете сите филмови на Вуди Ален (филмови кои ги режирал, филмови кои ги напишал и филмови во кои глумел).

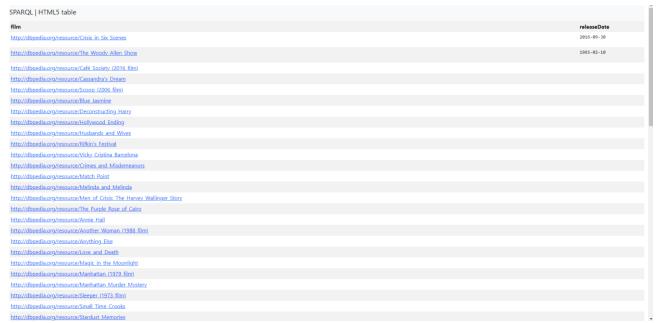




PREFIX dbr: http://dbpedia.org/ontology/>

4. Проширете го претходното SPARQL прашање за да ги добиете сите филмови на Вуди Ален, но подредени од најновите кон најстарите (според датумот на објавување).



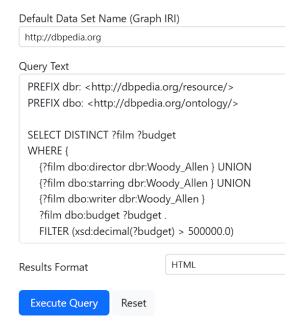


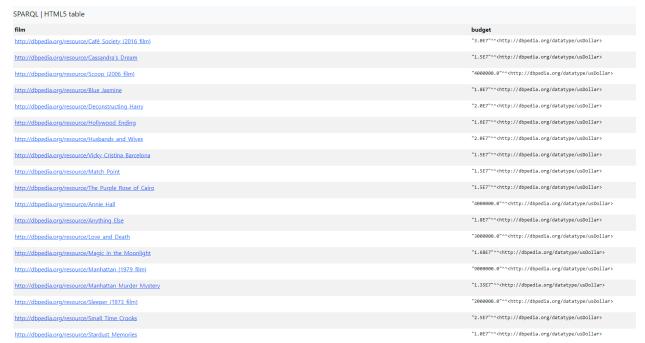
PREFIX dbr: PREFIX dbo: http://dbpedia.org/ontology/

SELECT DISTINCT ?film ?releaseDate WHERE {

{?film dbo:director dbr:Woody_Allen } UNION
{?film dbo:starring dbr:Woody_Allen } UNION
{?film dbo:writer dbr:Woody_Allen }
OPTIONAL {?film dbo:releaseDate ?releaseDate}
} ORDER BY DESC(?releaseDate)

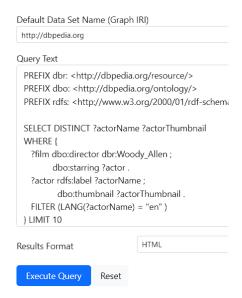
5. Проширете го претходното SPARQL прашање за да ги добиете сите филмови на Вуди Ален, со буџет поголем од \$500.000 долари.





PREFIX dbr: http://dbpedia.org/ontology/>

6. Напишете SPARQL прашање со кое ќе ги добиете имињата (label, на англиски) и сликите (thumbnail) на сите актери кои играле (starring) во филм на Вуди Ален. Ограничете го излезот на само 10 актери. Отстранете ги дупликатите од резултатите.



SPARQL HTML5 table	
actorName	actorThumbnail
"Carla Bruni"@en	http://commons.wikimedia.org/wiki/Special:FilePath/Carla_Bruni-Sarkozy_(3),jpg?width=300
"Carol Kane"@en	http://commons.wikimedia.org/wiki/Special:FilePath/Carol_Kane_(2018)_(cropped),jpg?width=300
"Carrie Fisher"@en	http://commons.wikimedia.org/wiki/Special:FilePath/Carrie_Fisher_2013-a_straightened.jpg?width=300
"Cate Blanchett"@en	http://commons.wikimedia.org/wiki/Special:FilePath/Cate_Blanchett_Cannes_2018_2_(cropped).jpg?width=300
"Roberto Benigni"@en	http://commons.wikimedia.org/wiki/Special:FilePath/Roberto_Benigni-5274.jpg?width=300
"Robin Williams"@en	http://commons.wikimedia.org/wiki/Special:FilePath/Robin_Williams_Happy_Feet_premiere.jpg?width=300
"Rosanna Arquette"@en	http://commons.wikimedia.org/wiki/Special:FilePath/Rosanna_ArquetteMonte-Carlo_Television_Festival.jpg?width=300
"Sam Waterston"@en	http://commons.wikimedia.org/wiki/Special:FilePath/Sam_Waterston_at_PaleyFest_2013.jpg?width=300
"Samantha Morton"@en	http://commons.wikimedia.org/wiki/Special:FilePath/Samantha_Morton_Edinburgh_International_Film_Festival.png?width=300
"Sandy Dennis"@en	http://commons.wikimedia.org/wiki/Special:FilePath/Sandy_Dennis.jpg?width=300

```
PREFIX dbr: <a href="http://dbpedia.org/resource/">PREFIX dbo: <a href="http://dbpedia.org/ontology/">PREFIX rdfs: <a href="http://www.w3.org/2000/01/rdf-schema#">PREFIX rdfs: <a href="http://www.w3.org/2000/01/rdf-schema#">PREFIX rdfs: <a href="http://www.w3.org/2000/01/rdf-schema#">PREFIX rdfs: <a href="https://www.w3.org/2000/01/rdf-schema#">PREFIX rdfs: <a href="https://www.w3.org/2000/01/rdf-schema#">https://www.w3.org/2000/01/rdf-schema#</a></a></a>

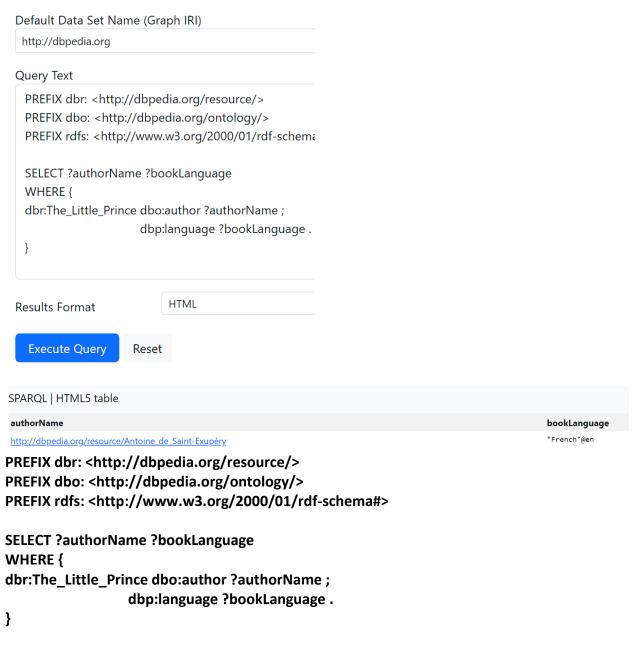
SELECT DISTINCT ?actorName ?actorThumbnail

WHERE {
    ?film dbo:director dbr:Woody_Allen;
    dbo:starring ?actor.
    ?actor rdfs:label ?actorName;
    dbo:thumbnail ?actorThumbnail .

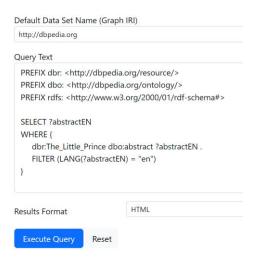
FILTER (LANG(?actorName) = "en")
} LIMIT 10
```

II. Работа со податоци за книга

7. Отворете го ресурсот http://dbpedia.org/resource/The_Little_Prince. Погледнете ги својствата (тројките) кои ги има оваа книга. Напишете SPARQL прашање со кое ќе ги добиете авторот и јазикот на книгата.



8. Напишете SPARQL прашање со кое ќе го добиете апстрактот на овој ресурс. Апстрактот нека биде на англиски јазик.



```
SPARQL | HTML5 table

abstractEN

"The Little Prince (French: Le Petit Prince, pronounced [la p(a)ti pad5]) is a novella by French aristocrat, writer, and military aviator Antoine de Saint-Exupéry. It was first published in English

PREFIX dbr: <a href="http://dbpedia.org/resource/">http://dbpedia.org/resource/</a>

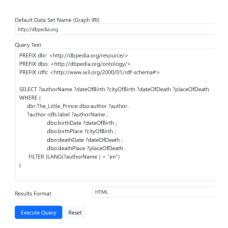
PREFIX dbo: <a href="http://dbpedia.org/ontology/">http://dbpedia.org/ontology/</a>

PREFIX rdfs: <a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#>

SELECT ?abstractEN

WHERE {
    dbr:The_Little_Prince dbo:abstract ?abstractEN .
    FILTER (LANG(?abstractEN) = "en")
}
```

9. Преку истиот ресурс (книгата "Малиот Принц"), освен името на авторот извлечете ги информациите и за датумот и местото на раѓање на авторот, како и датумот и местото на умирање.



SPARQL HTML5 table				
authorName	dateOfBirth	cityOfBirth	dateOfDeath	placeOfDeath
"Antoine de Saint-Exupéry"@en	1900-06-29	http://dbpedia.org/resource/Lyon	1944-07-31	http://dbpedia.org/resource/Marseille
"Antoine de Saint-Exupéry"@en	1900-06-29	http://dbpedia.org/resource/Lyon	1944-07-31	http://dbpedia.org/resource/Mediterranean_Sea
"Antoine de Saint-Exupéry"@en	1900-06-29	http://dbpedia.org/resource/Lyon	1944-07-31	http://dbpedia.org/resource/Occupied France

PREFIX dbr: http://dbpedia.org/ontology/>

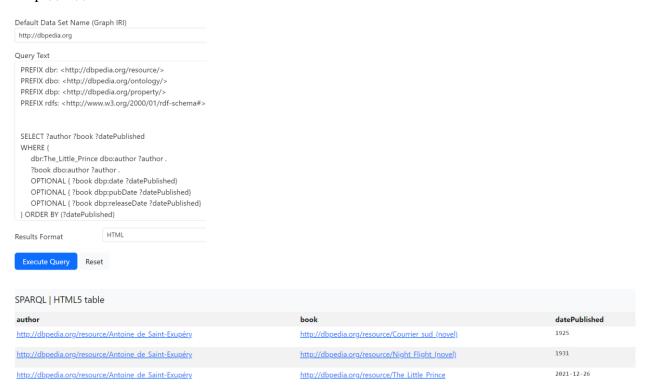
PREFIX rdfs: http://www.w3.org/2000/01/rdf-schema#>

SELECT ?authorName ?dateOfBirth ?cityOfBirth ?dateOfDeath ?placeOfDeath WHERE {

```
dbr:The_Little_Prince dbo:author ?author .
?author rdfs:label ?authorName;
dbo:birthDate ?dateOfBirth;
dbo:birthPlace ?cityOfBirth;
dbo:deathDate ?dateOfDeath;
dbo:deathPlace ?placeOfDeath .
FILTER (LANG(?authorName) = "en")
```

}

10. Преку истиот ресурс (книгата "Малиот Принц"), извлечете ги имињата и на останатите книги кои ги напишал нејзиниот автор, нивните години на издавање и сортирајте ги според годината.



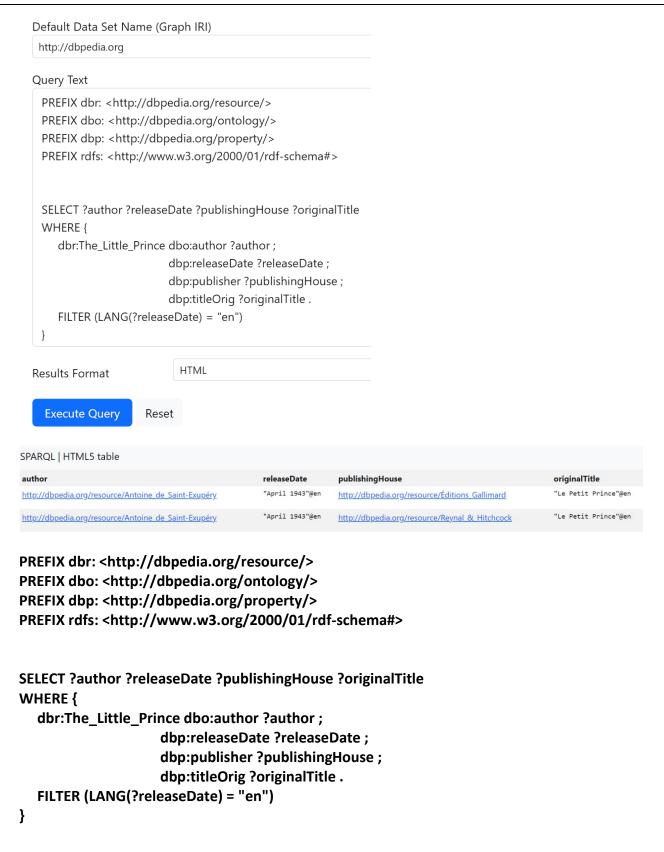
PREFIX dbr: http://dbpedia.org/ontology/>PREFIX dbp: http://dbpedia.org/property/>

} ORDER BY (?datePublished)

PREFIX rdfs: http://www.w3.org/2000/01/rdf-schema#>

SELECT ?author ?book ?datePublished WHERE { dbr:The_Little_Prince dbo:author ?author . ?book dbo:author ?author . OPTIONAL { ?book dbp:date ?datePublished} OPTIONAL { ?book dbp:pubDate ?datePublished} OPTIONAL { ?book dbp:releaseDate ?datePublished}

11. Напишете SPARQL прашање за да добиете одговор во која година прв пат е издадена оваа книга, од која издавачка куќа и под кој оригинален наслов.

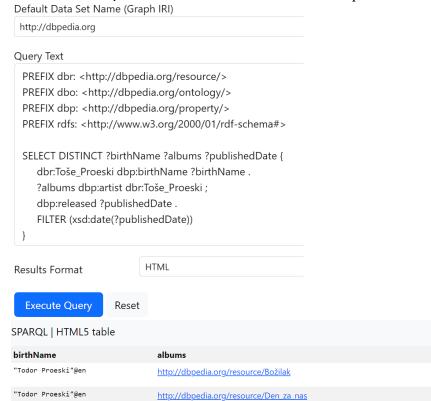


III. Работа со податоци за ресурс по слободен избор

12. Отворете Wikipedia article по ваш избор. Отворете го соодветниот DBpedia ресурс и разгледајте ги тројките кои постојат за тој ресурс. Напишете барем три SPARQL прашања по ваш избор, во кој ќе ги пробате останатите опции од аудиториски вежби.

Пример идеи:

- Музички бенд / артист:
 - о прикажете ги имињата и датумите на објава на албумите;
 - о прикажете ги имињата и местата на раѓање на членовите на бендот;



http://dbpedia.org/resource/Sinot_božji

http://dbpedia.org/resource/Po_Tebe_(song)

http://dbpedia.org/resource/Po_tebe_(album)

http://dbpedia.org/resource/Po_tebe_(album)

http://dbpedia.org/resource/Ako me pogledneš vo oči

http://dbpedia.org/resource/The_Hardest_Thing_(album)

PREFIX dbr:
PREFIX dbo:
PREFIX dbp:
prefix dbp: <a hre

PREFIX rdfs: http://www.w3.org/2000/01/rdf-schema#>

```
SELECT DISTINCT ?birthName ?albums ?publishedDate {
    dbr:Toše_Proeski dbp:birthName ?birthName .
    ?albums dbp:artist dbr:Toše_Proeski ;
    dbp:released ?publishedDate .
    FILTER (xsd:date(?publishedDate))
}
```

• Личност:

"Todor Proeski"@en

"Todor Proeski"@en

"Todor Proeski"@en

"Todor Proeski"@en

"Todor Proeski"@en

"Todor Proeski"@en

- о прикажете краток опис и биографија, на англиски;
- о прикажете ги датумите и местата на раѓање / смрт;
- о прикажете ги линковите до неговиот / нејзиниот веб сајт, социјален профил, итн.;

publishedDate

2006

2004

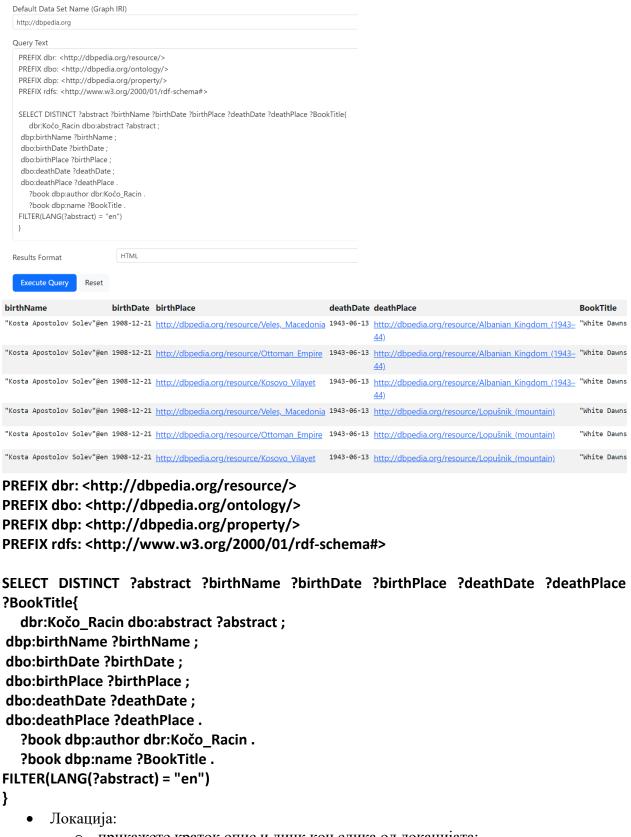
2000

2005

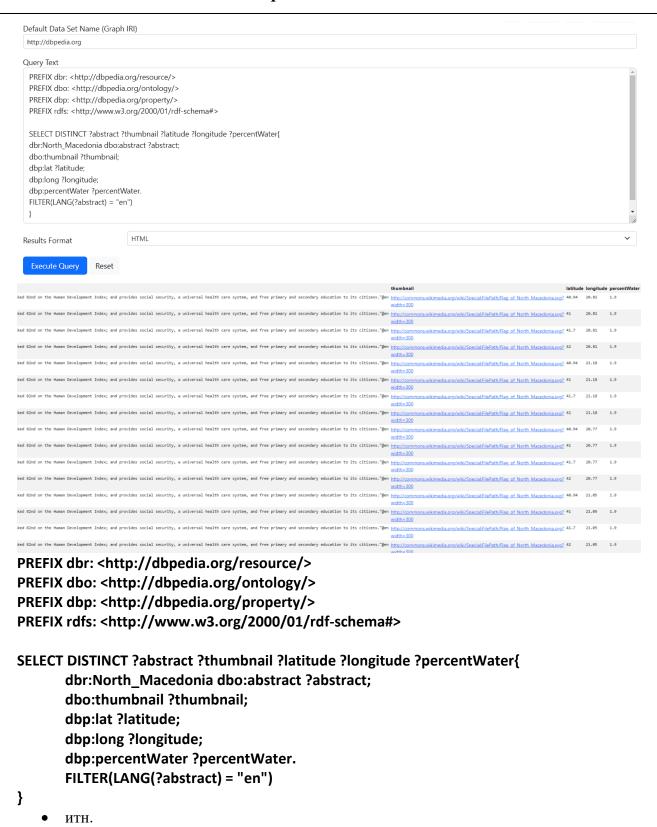
""@en

2002-10-25

2009-01-25



- прикажете краток опис и линк кон слика од локацијата;
- прикажете ги longitude и latitude координатите, како и надморската височина;



IV. Работа со податоци за потребите на конкретна апликација

13. Размислете за некоја мобилна или веб апликација која би сакале да ја развивате. Согласно темата, пробајте да напишете едно или повеќе SPARQL прашања со кои од DBpedia ќе ги извлечете сите потребни податоци за нејзино реализирање.

Пример идеи:

 најдете ги сите филмови објавени во тековната година и прикажете ги нивните имиња, краток опис, жанр, како и датумот на објава; сортирајте ги по датум на објава;

```
Default Data Set Name (Graph IRI)
   http://dbpedia.org
 Query Text
     PREFIX dbo: <a href="http://dbpedia.org/ontology/">http://dbpedia.org/ontology/>
     PREFIX rdfs: <a href="http://www.w3.org/2000/01/rdf-schema">http://www.w3.org/2000/01/rdf-schema#>
    PREFIX xsd: <a href="http://www.w3.org/2001/XMLSchema#">http://www.w3.org/2001/XMLSchema#</a>
    SELECT DISTINCT ?movie ?name ?releaseDate ?productionCompany ?abstract
    WHERE {
      ?movie a dbo:Film;
             dbo:abstract ?abstract;
             rdfs:label ?name;
             dbo:releaseDate ?releaseDate ;
             dbo:productionCompany?productionCompany.
     FILTER(YEAR(NOW()) = (YEAR(?releaseDate)))
     FILTER(LANG(?name) = "en")
     FILTER(LANG(?abstract) = "en")
     ORDER BY ?releaseDate LIMIT 25
                                                     HTML
  Results Format
     Execute Query
                                    Reset
SPARQL | HTML5 table
      bpedia.org/resource/The Old Way
                                                                                     2023-01-06 http://dbpedia.org/resource/Saturn_Films
http://dbpedia.org/resource/Teen_Wolf:_The_Movie
                                                   "Teen Wolf: The Movie"@en
                                                                                    2023-01-26 http://dbpedia.org/resource/MGM_Television
                                                                                                                                               "Teen Wolf: The Movie is an upcoming American supernatural
    //dbpedia.org/resource/Teen_Wolf:_The_Movie
                                                    "Teen Wolf: The Movie"@en
                                                                                     2023-01-26 <a href="http://dbpedia.org/resource/MTV_Entertainment_Studios">http://dbpedia.org/resource/MTV_Entertainment_Studios</a> "Teen Wolf: The Movie is an upcoming American supernatural
                                                   "Teen Wolf: The Movie"@en
                                                                                  2023-01-26 http://dbpedia.org/resource/Orion_Pictures "Teen Wolf: The Movie is an upcoming American supernatural
http://dbpedia.org/resource/Teen_Wolf:_The_Movie
http://dbpedia.org/resource/Selfiee (2023 film)
                                                    "Selfiee (2023 film)"@en
                                                                                     2023-02-24 http://dbpedia.org/resource/Dharma_Productions
                                                                                                                                                "Selfiee is a 2023 Indian Hindi-language comedy-drama film
                                                    "Selfiee (2023 film)"@en
                                                                                                                                               "Selfiee is a 2023 Indian Hindi-language comedy-drama film
http://dbpedia.org/resource/Selfiee_(2023_film)
                                                                                    2023-02-24 http://dbpedia.org/resource/Cape_of_Good_Films
                                                    "Selfiee (2023 film)"@en
http://dbpedia.org/resource/Selfiee_(2023_film)
                                                                                     2023-02-24 http://dbpedia.org/resource/Magic_Frames
http://dbpedia.org/resource/Selfiee_(2023_film)
                                                   "Selfiee (2023 film)"@en
                                                                                                                                               "Selfiee is a 2023 Indian Hindi-language comedy-drama film of
                                                                                    2023-02-24 http://dbpedia.org/resource/Prithviraj_Productions
    //dbpedia.org/resource/Money_Back_Guarantee_(2023_film) "Money_Back_Guarantee (2023_film)"@en_2023-04-21 http://dbpedia.org/resource/Zashko_Films
                                                                                                                                                "Money Back Guarantee is a Pakistani film directed & writte
http://dbpedia.org/resource/The_Nun_2 "The Nun_2"@en
                                                                             2023-09-08 http://dbpedia.org/resource/The_Safran_Company.
                                                                                     2023-09-08 http://dbpedia.org/resource/Atomic Monster Productions "The Nun 2 is an upcoming American supernatural horror film
http://dbpedia.org/resource/The Nun 2
                                                    "The Nun 2"@en
http://dbpedia.org/resource/The_Nun_2
                                                                                     2023-09-08 <a href="http://dbpedia.org/resource/New Line Cinema">http://dbpedia.org/resource/New Line Cinema</a> "The Nun 2 is an upcoming American supernatural horror film
```

PREFIX dbo: http://dbpedia.org/ontology/>

PREFIX rdfs:
PREFIX xsd: http://www.w3.org/2001/XMLSchema">

SELECT DISTINCT ?movie ?name ?releaseDate ?productionCompany ?abstract WHERE {

```
?movie a dbo:Film;
dbo:abstract ?abstract;
rdfs:label ?name;
dbo:releaseDate ?releaseDate;
```

```
dbo:productionCompany ?productionCompany .

FILTER(YEAR(NOW()) = (YEAR(?releaseDate)))

FILTER(LANG(?name) = "en")

FILTER(LANG(?abstract) = "en")

}

ORDER BY ?releaseDate LIMIT 25
```

- најдете ги музичките албуми кои се објавени во тековната година и прикажете го артистот, името на албумот, жанрот, датумот на објава и линкот кон слика од албумот; сортирајте ги по датум на објава;
- итн.