Домашно бр. 2 SPARQL

А) Домашна задача

1. IIITO e SPARQL?

SPARQL е прашален јазик за извлекување податоци од Семантичкиот Веб. За РДФ, овој јазик е исто што и SQL за релациони бази на податоци.

2. Кои типови на SPARQL прашања постојат?

SELECT, DESCRIBE, ASK, UPDATE, CONSTRUCT.

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

SUBJECT-PREDICATE-OBJECT

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

Со користење на **OPTIONAL** блокот и наведување на патернот во него.

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

SPARQL endpoints се креирани со цел да може да им се пристапува преку GET, POST барања. Преку ова, тие се дел од RESTful сервисите, што ги прави поволни за користење од страна на апликации! Ова е голема предност за разлика од пристапот со АПИ, каде при секоја промена, би требало да се конфигурира секоја апликација што го користи истото АПИ.

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:
 - а. 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/
 - c. 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.

Query:

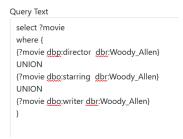


Response:



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).

Query:

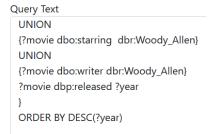


Response:



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).

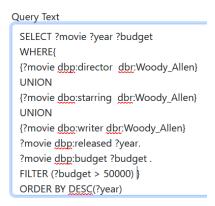
Query:



Response:

movie	year
http://dbpedia.org/resource/Antz	1998-11-03
http://dbpedia.org/resource/Don't_Drink_the_Water_(1994_film)	1994-12-18
http://dbpedia.org/resource/Meetin'_WA	1986
http://dbpedia.org/resource/Casino_Royale_(1967_film)	1967

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



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.

Query:

Response:



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.

Query:

```
query Text

select ?author ?lang
where {
dbr:The_Little_Prince dbo:author ?author;
dbp:language ?lang
}
```

Response:



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

Query:

```
Query Text
select ?abstract
where {
  dbr:The_Little_Prince dbo:abstract ?abstract.
FILTER(lang(?abstract)="en")
}
```

Response:

```
SPARQL | HTML5 table

abstract
```

"The Little Prince (French: Le Petit Prince, pronounced [lə p(ə)ti pušs]) is a novella by French aristocrat, writer, and military aviator Antoine de Saint-Exupéry.

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.

Query:

```
query Text

select ?birthplace ?birthdate ?deathplace ?deathdate where {
    dbr:The_Little_Prince dbo:author ?author.
    ?author dbo:birthPlace ?birthplace;
    dbo:birthDate ?birthdate;
    dbo:deathPlace ?deathplace;
    dbo:deathDate ?deathdate
}
```

Response:

SPARQL HTML5 t	table	
birthplace	birthdate deathplace	deathdate
http://dbpedia.org /resource/Lyon	1900-06-29 http://dbpedia.org/resource /German_military_administration_in_occupied_France_during_World_War_II	1944-07-31
http://dbpedia.org /resource/Lyon	1900-06-29 http://dbpedia.org/resource/Marseille	1944-07-31
http://dbpedia.org /resource/Lyon	1900-06-29 http://dbpedia.org/resource/Mediterranean_Sea	1944-07-31

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.

Query:

```
Query Text

SELECT ?book ?releaseDate

WHERE{
    dbr:The_Little_Prince dbo:author ?author.
    ?book dbo:author ?author.
    ?book dbp:releaseDate ?releaseDate
    }

ORDER BY ?releaseDate
```

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.

Query:

```
Query Text

SELECT ?releaseDate ?publisher ?originalTitle

WHERE{

dbr:The_Little_Prince dbp:releaseDate ?releaseDate;

dbo:publisher ?publisher;

dbp:titleOrig ?originalTitle

}
```

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;
- Личност:
 - 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.;
- Локација:
 - о прикажете краток опис и линк кон слика од локацијата;
 - о прикажете ги longitude и latitude координатите, како и надморската височина;
- итн.

Прикажување на информации за актерот Том Харди. Абстрактот, краток опис, датум на раѓање, место на раѓање и линкови кои водат до ресурси за него бидејќи не се достапни неговите социјални мрежи.

```
Query Text

SELECT ?abstract ?shortDesc ?bday ?birthplace ?links

WHERE{
    dbo:Tom_Hardy dbo:abstract ?abstract;
        rdfs:comment ?shortDesc;
    dbo:birthDate ?bday;
    dbo:birthPlace ?birthplace;
    dbo:wikiPageExternalLink ?links.

FILTER(lang(?abstract)="en")
}
```

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;
- 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.

Пребарување на сите песни издадени во 2022 година. Приказ на нивните артисти, жанр, датум на издавање и линк кон слика. Сортирани според датумот на издавање. Со филтер командата се пребарува УРИ-то за 2022 песни без точно да го наведеме, со помош на регекс.

```
SELECT ?song ?artist ?genre ?releasedate ?thumbnail WHERE{
?concept a skos:Concept.
?concept rdfs:label ?y .
?song dbo:wikiPageWikiLink ?concept;
    dbp:artist ?artist;
    dbo:genre ?genre;
    dbo:releaseDate ?releasedate;
    dbo:thumbnail ?thumbnail.
FILTER regex(?y, "2022 songs", "i")
}
ORDER BY ?releaseDate
```

song	artist	genre	releasedate	thumbnail
http://dbpedia.org/resource /Don't Wake Me_Up_(Jonas_Blue_and_Why_Don't_We_song)	"Jonas Blue and Why Don't We"@en	http://dbpedia.org/resource /Dance-pop	2022-01-07	http://commons.wikimedia.org/wiki/Special:FilePa /Jonas Blue - Don't Wake Me Up.jpg?width=300
http://dbpedia.org/resource/Sentimentai	http://dbpedia.org/resource/Monika_Liu	http://dbpedia.org/resource /Pop_music	2022-01-18	http://commons.wikimedia.org/wiki/Special:FilePa/Sentimentai.jpg?width=300
http://dbpedia.org/resource/Same_God	"Elevation Worship featuring Jonsal Barrientes"@en	http://dbpedia.org/resource /Contemporary_worship_music	2022-01-21	http://commons.wikimedia.org/wiki/Special:FilePa /Same_God_(feat_Jonsal_Barrientes) _ Elevation_Worship.jpg?width=300
http://dbpedia.org/resource/Emo_Girl	"Machine Gun Kelly and Willow"@en	http://dbpedia.org/resource /Pop-punk	2022-02-04	http://commons.wikimedia.org/wiki/Special:FilePa /Machine Gun Kelly Willow Emo Girl.png?width=
http://dbpedia.org/resource/Saoko	http://dbpedia.org/resource/Rosalía_(singer)	http://dbpedia.org/resource /Alternative_reggaeton	2022-02-04	http://commons.wikimedia.org/wiki/Special:FilePa/Rosalia - Saoko.png?width=300
http://dbpedia.org/resource/Saoko	http://dbpedia.org/resource/Rosalía_(singer)	http://dbpedia.org/resource /Experimental_music	2022-02-04	http://commons.wikimedia.org/wiki/Special:FilePa/Rosalía - Saoko.png?width=300
http://dbpedia.org/resource/Pow_(Elvana_Gjata_song)	http://dbpedia.org/resource/Elvana_Gjata	http://dbpedia.org/resource /Contemporary_R&B	2022-02-11	http://commons.wikimedia.org/wiki/Special:FilePa /York Street_station_NYC_Subway.jpg?width=300
http://dbpedia.org/resource/Pow_(Elvana_Gjata_song)	http://dbpedia.org/resource/Elvana_Gjata	http://dbpedia.org/resource /Urban_contemporary_music	2022-02-11	http://commons.wikimedia.org/wiki/Special:FilePa /York Street station NYC Subway.jpg?width=300
http://dbpedia.org/resource/Mud_Mud_Ke	"Tony Kakkar and Neha Kakkar"@en	http://dbpedia.org/resource /Indian_pop	2022-02-12	http://commons.wikimedia.org/wiki/Special:FilePa /Mud_Mud_Ke_cover_art.jpg?width=300
http://dbpedia.org/resource/Mr. (Yoasobi_song)	http://dbpedia.org/resource/Yoasobi	http://dbpedia.org/resource /City_pop	2022-02-16	http://commons.wikimedia.org/wiki/Special:FilePa /Yoasobi - Mr.jpg?width=300