

Interrogazione Repository semantici con SPARQL

SPARQL, linguaggio di interrogazione dell' RDF

<https://www.w3.org/TR/rdf-sparql-query/>

SPARQL è un linguaggio di interrogazione RDF (Resource Description Framework) uno standard W3C per la codifica,e lo scambio di metadati strutturati che consente l'interoperabilità semantica fra metadati.

SPARQL è molto simile a SQL e adotta la sintassi TURLE, un'estensione di N-Triples, alternativa estremamente sintetica e intuitiva al tradizionale RDF/XML.

Le query SPARQL si basano sul meccanismo del "pattern matching" e in particolare su un costrutto, il "triple pattern": **Soggetto** → **Predicato** → **Oggetto**

```
PREFIX cd: <http://example.org/cd/>
SELECT ?titolo ?autore ?anno
FROM <http://cd.com/listacd.ttl>
WHERE {?titolo cd:autore ?autore.
      ?titolo cd:anno ?ann .
}
```

PREFIX definisce i namespace

SELECT definisce le variabili che devono essere restituite

FROM definisce il repository

WHERE definisce il “triple pattern matching”

?titolo cd:autore ?autore.

Interrogo tutto le risorse che abbiano un autore

?resource edm:type 'VIDEO'

Interrogo tutto le risorse il cui tipo sia VIDEO.

Dove ?resource è il Soggetto, edm:type il predicato e VIDEO il soggetto

Posso aggiungere inoltre dei filtri sia su dati numerici che su dati testuali. Per i testi i filtro funziona utilizzando le espressioni regolari REGEX

FILTER (?anno > 2000)

FILTER regex(?title , '^libro','i'). }

Filtro le risorse il cui titolo inizia per “libro” non case sensitive

posso definire dei valori opzionali

```
OPTIONAL {?titolo cd:anno ?anno}
```

e ordinare e limitare i valori

```
ORDER BY DESC(?autore)  
LIMIT 10  
OFFSET 10
```

SPASQL permette anche unire i risultati provenienti da repository differenti tramite UNION

Da approfondire

Europeana data Model

<http://pro.europeana.eu/page/edm-documentation>

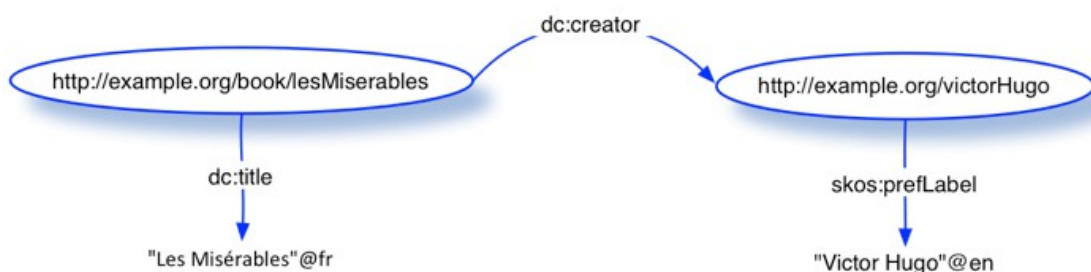
Europeana è un repository che si pone come catalogo per la cultura europea costruendo/definendo un modello semantico per la catalogazione semantica dei dati EDM (European Data Model)

Europeana dispone di un ENPOINT SPARQL per interrogazione RDF

<http://labs.europeana.eu/api/linked-open-data-sparql-endpoint>

Si basa principalmente sui seguenti standard:

- OAI ORE (Open Archives Initiative Object Reuse & Exchange) for organizing an object's metadata and digital representation(s)
- Dublin Core for descriptive metadata
- SKOS (Simple Knowledge Organization System) for conceptual vocabulary representation
- CIDOC-CRM for event and relationships between objects



Applicazione

PHP7, HTML5

Librerie usate

ARC2

<https://github.com/semsol/arc2/wiki>

Per gestire interrogazioni SPARSQL

SLIM FRAMEWORK

<https://www.slimframework.com/>

Per gestire le rotte e i template e il engine TWIG

Struttura applicazione

public

Contiene la parte pubblica e principale della applicazione il file index.php definisce tutta la logica della app chiamando la rotta /playlist e la gestione delle variabili di sessione

src

Contiene le librerie usate ARC2 e una classe costruita ad hoc **sparqlQuery** per effettuare le interrogazioni ai repository utilizzando un metodo statico

```
sparqlQuery::fetch($endpoint, $query);
```

templates

contiene i template html e javascript

vendor

contiene tutte le librerie e i componenti slim framework

QUERY

Le Query effettuata è la seguente

```
FILTER  
PREFIX edm: <http://www.europeana.eu/schemas/edm/> PREFIX ore:  
<http://www.openarchives.org/ore/terms/> PREFIX dc:  
<http://purl.org/dc/elements/1.1/> SELECT *  
WHERE {  
    ?resource edm:type 'IMAGE' ;  
    ore:proxyIn ?proxy ;  
    dc:title ?title ;  
    dc:description ?description ;  
    dc:creator ?creator ;  
    dc:source ?source ; dc:date  
    ?date. ?proxy edm:isShownBy  
    ?mediaURL .  
    FILTER regex(?title , '^','i').  
}  
OFFSET 100  
LIMIT 5
```

```
?resource edm:type 'IMAGE' ;
```

Il tipo viene impostato dinamicamente in base dalla select

```
FILTER regex(?title , '^','i').
```

Il titolo impostato in base al campo di input, questa ricerca i titoli che iniziano con la stringa inserita nel campo input, sia maiuscoli che minuscoli.

```
FILTER (?date > '2000' && ?date < '2016')
```

filtra risorse con data compresa fra 2000 2016

LIMIT è posto a 5

OFFSET invece viene generato casualmente per far tornare diversi contenuti, questo comporta però che a volte il filtro per titolo non funziona.

Conclusioni

L'applicativo rappresenta una prima versione basilare per l'interrogazione di repository RDF con SPASQL.

Rimane ancora da implementare la parte di interrogazione di più repository e approfondire sia il linguaggio SPARQL che i modelli rappresentativo dei varir epository EDM, DBPEDIA ecc..

Migliorare il filtro su mediaURL per far tornare solo risorse mp4 o mp3

Alcune volte i player non funzionano perché ritornano video mpg o audio non compatibili.

Sarebbe da perfezionare il filtro chiedendo solo di ritornare audio mp3 o video mpg4.

Rimane sa migliorare tutta la parte frontend sia dal punto di vista di usabilità che grafica che con l'introduzione di chiamate ajax.

