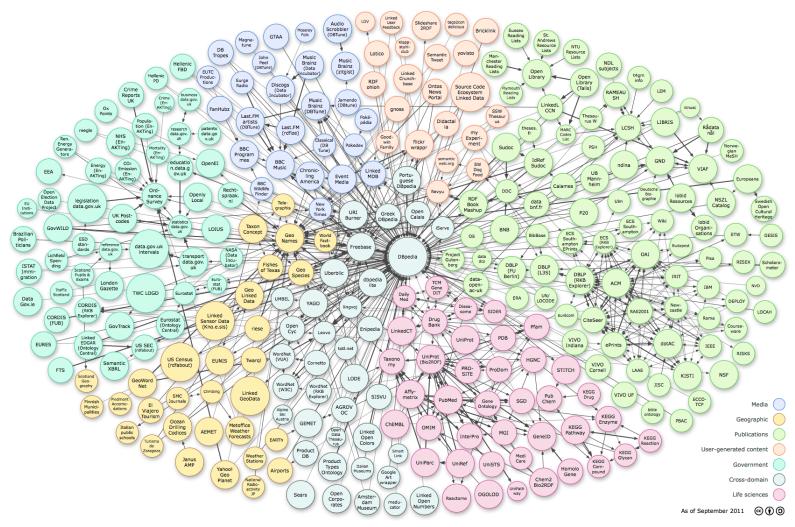
Module MLBDA Master Informatique Spécialité DAC

Cours 9 – SPARQL

Linking Open Data

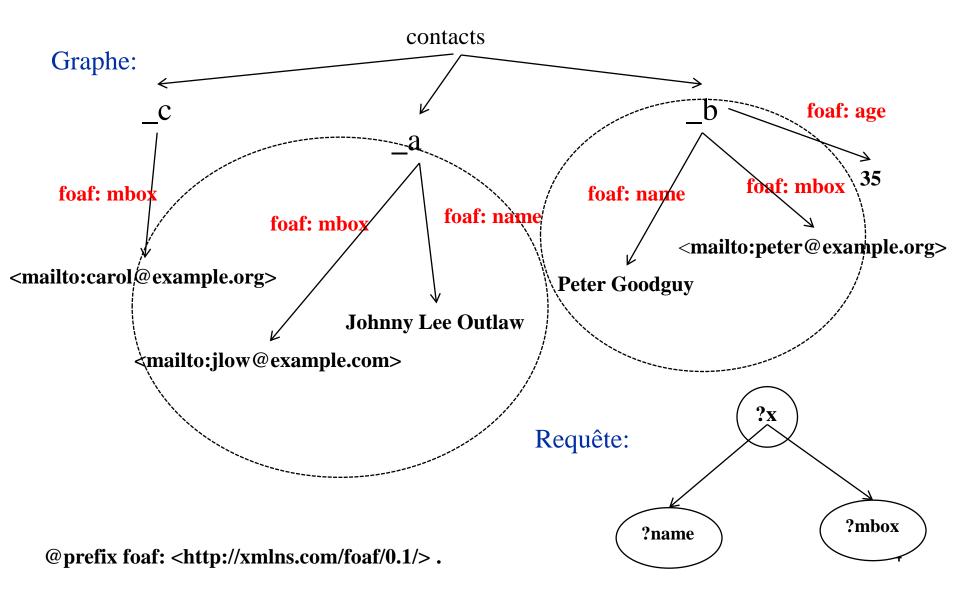


Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. http://lod-cloud.net/

SPARQL

- SPARQL: Simple Protocol and RDF Query Language
- Langage de requêtes du W3C pour RDF/RDFS
 - SPARQL 1.0 : recommandation 2008
 - SPARQL 1.1 : recommandation 2013
- Principe : pattern matching sur des graphes
- Plusieurs formes de requêtes
 - Select
 - Construct, Describe
 - Ask

Exemple



Exemple

Données (un graphe):

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .

_:a foaf:name "Johnny Lee Outlaw" .

_:a foaf:mbox <mailto:jlow@example.com> .

_:b foaf:name "Peter Goodguy" .

_:b foaf:mbox <mailto:peter@example.org> .

_:b foaf:age 35.

_:c foaf:mbox <mailto:carol@example.org> .
```

Requête de base

- PREFIX : permet de déclarer les espaces de noms.
- **SELECT**: permet d'identifier les variables du résultat (parmi les variables de la clause WHERE).
- WHERE définit le motif de graphe recherché à l'aide de triplets (comme en RDF).

Exemple

Réponse:

Name	Mbox
"Johnny Lee Outlaw"	<pre><mailto:jlow@example.com></mailto:jlow@example.com></pre>
"Peter Goodguy"	<pre><mailto:peter@example.org></mailto:peter@example.org></pre>

Littéraux et nœuds blancs

• Les littéraux : chaînes de caractères entre quotes.

```
"valeur", @motcle, ^^type,

"false"^^xsd:boolean, "1.3"^^xsd:decimal
```

- On peut simplifier et écrire
 - false au lieu de "false"^^xsd:boolean
 - 1.3 au lieu de "1.3" \ xsd:decimal
- Les nœuds blancs sont désignés par un label (_xx) ou par []

```
[ :a "toto" ] :b "titi" .
est équivalent à
    _xx :a "toto" .
    xx :b "titi" .
```

Espaces de noms

• On peut définir un préfixe associé à un espace de noms, et un espace de nom "base" de la requête.

```
PREFIX dc: <a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/>
SELECT ?title
WHERE { <http://example.org/book/book1> dc:title ?title }
PREFIX dc: <a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/>
PREFIX : <http://example.org/book/>
SELECT Stitle
WHERE { :book1 dc:title $title }
BASE <a href="http://example.org/book/">
PREFIX dc: <a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/>
SELECT $title
WHERE { <book1> dc:title $title }
```

Motif de triplets

```
L'ensemble des termes RDF est défini par T = U \cup L \cup B
   U : ensemble des URI
   <http://www.w3.org/TR/rdf-syntax-grammar>
  dc:title
   L: littéraux RDF (valeurs): "valeur", @motcle, ^^type
    "Dave Beckett", @prefix, "false"^^xsd:boolean,
   "1.3"^^xsd:decimal
   B: nœuds blancs
Un triplet RDF est un élément de l'ensemble (U ∪ B) U
Un motif de triplets est un élément de l'ensemble (T U V) (U U V) (T U V)
   où V est un ensemble de variables tel que V \cap T = \emptyset
Les variables sont préfixées par ? ou par $
```

Exemples

- Motif ∈ V U V
 ?book dc:title ?title .
- Motif ∈ B U L (avec définition de préfixe)

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
_:a foaf:name "Bob" .
```

Motifs de triplets

Pour simplifier l'écriture des motifs de triplets, on peut

 Les factoriser ?x foaf:name ?name; foaf:mbox ?mbox . ?x foaf:nick "Robert", "Bob" Utiliser les collections RDF (1 ?x 3) :a "toto".équivaut à :b1 rdf:first 1; rdf:rest :b2 . _:b2 rdf:first ?x; rdf:rest _:b3 . _:b3 ref:first 3; rdf:rest rdf:nil . :b1 :a "toto" . Utiliser des abréviations (rdf:type) ?x a :Class1 équivaut à ?x rdf:type :Class1

Motifs de graphes

- Un ensemble de *motifs de triplets* est un motif de graphe (élémentaire)
- Si GP est un motif de graphe, {GP} est un motif de graphe de groupe.
- Si GP et GP' sont des motifs de graphe :
 - GP FILTER (test) GP' : sélection/filtrage
 - GP OPTIONAL GP': graphes optionnels
 - GP UNION GP': union de graphes
 sont des motifs de graphe.

Exemples de motifs de graphe

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
• Motif de graphe élémentaire :
      _:a foaf:age 26 . _:a foaf:mbox ?m .
• Motif de groupe :
      { _:a foaf:age 26 . _:a foaf:mbox ?m . }
• Motif de groupe avec FILTRE :
      { _:a foaf:age ?age .
      FILTER (?age < 26)
      _:a foaf:mbox ?m . }
• Motif de groupe avec OPTIONAL :
{ _:a foaf:age 26 . OPTIONAL { _:a foaf:mbox ?m . }}
```

Filtres

• Le mot-clef **FILTER** permet de restreindre les solutions sur tout le groupe où le filtre apparaît. La position du filtre dans le groupe n'a pas d'importance, les motifs suivants sont équivalents :

```
{ ?x ns:price ?price .
FILTER (?price < 15)
?x dc:title ?title . }
{FILTER (?price < 15)
?x ns:price ?price .
?x dc:title ?title . }
{ ?x ns:price ?price .
?x dc:title ?title .
FILTER (?price < 15) }
```

Motif Optionnel

- Un motif de graphe élémentaire permet de rechercher des solutions qui correspondent entièrement au motif d'interrogation.
- Le filtrage optionnel (mot-clé **OPTIONAL**) permet d'obtenir des solutions même si des parties du motif d'interrogation ne correspondent pas.

Filtrage des alternatives

• Le mot-clef **UNION** permet d'indiquer des alternatives de motifs (un motif OU un autre peut correspondre). Si plusieurs alternatives correspondent, toutes les solutions sont trouvées.

Requête SPARQL

```
Une requête SPARQL est de la forme
       <FORMAT>
          FROM source
          WHERE { motif }
       <TRANSFORM>
FORMAT> définit le format du résultat : SELECT, CONSTRUCT,
  DESCRIBE, ASK
FROM définit la source RDF (optionnel si source par défaut)
motif est un motif de graphe
TRANSFORM> est un transformateur : order, limit, offset
```

Formats

format	résultat
SELECT, SELECT DISTINCT	table de données
CONSTRUCT	graphe RDF
ASK	valeur Booléenne (match non-vide)
DESCRIBE	description des ressources trouvées

Clause FROM

- La clause FROM est utilisée pour indiquer l'IRI d'une ressource sur lequel effectuer la requête.
- En l'absence de clause FROM, l'interrogation s'effectue sur le graphe par défaut.

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
:a foaf:name "Alice" foaf:mbox
  <mailto:alice@work.example> .
:b foaf:name "Bob" foaf:mbox
  <mailto:bob@work.example> .
Requête
PREFIX foaf: <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/>
SELECT ?name
FROM <http://example.org/foaf/aliceFoaf>
WHERE { ?x foaf:name ?name }
```

Sémantique

- Une solution de motif (de graphe ou de triplet) est une fonction de substitution d'un ensemble de variables vers l'ensemble des termes
- Une solution de motif S est une solution pour une requête **format> FROM source WHERE { motif }** si S appliqué au motif est un sous-graphe de **source.**
- Le résultat d'une requête est l'ensemble de toutes ses solutions de motifs.

```
Données:
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
    rdf:type
                    foaf:Person .
:a
    foaf:name
                    "Alice" .
:a
    foaf:mbox
                    <mailto:alice@example.com> .
:a
:a foaf:mbox
                    <mailto:alice@work.example> .
:b rdf:type
                    foaf:Person .
:b foaf:name
                    "Bob" .
```

Résultat :

name	mbox
"Alice"	<mailto:alice@example.com></mailto:alice@example.com>
"Alice"	<mailto:alice@work.example></mailto:alice@work.example>
"Bob"	

```
Données:
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
    rdf:type
                    foaf:Person .
:a
    foaf:name
                     "Alice" .
:a
    foaf:homepage
                    <http://work.example.org/alice/> .
:a
_:b rdf:type
                    foaf:Person .
:b foaf:name
                     "Bob" .
:b foaf:mbox
                     <mailto:bob@work.example> .
```

Résultat :

name	mbox	hpage
"Alice"		http://work.example.org/alice/>
"Bob"	<mailto:bob@work.example></mailto:bob@work.example>	

```
Requête: plusieurs motifs OPTIONAL (pas équivalent à la requête précédente)

PREFIX foaf: <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/>

SELECT ?name ?mbox ?hpage

WHERE { ?x foaf:name ?name .

OPTIONAL { ?x foaf:mbox ?mbox .

?x foaf:homepage ?hpage }

}
```

Résultat :

name	mbox	hpage
"Alice"		
"Bob"		

Données:

```
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix : <http://example.org/book/> .
@prefix ns: <http://example.org/ns#> .
:book1 dc:title "SPARQL Tutorial" .
:book1 ns:price 42 .
:book2 dc:title "The Semantic Web" .
:book2 ns:price 23 .
```

```
Requête:filtrage
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX ns: <http://example.org/ns#>
SELECT ?title ?price
WHERE { ?x ns:price ?price .
FILTER (?price < 30) .
?x dc:title ?title . }</pre>
```

Résultat :

Title	Price
"The Semantic Web"	23

Résultat:

Title	Price
"SPARQL Tutorial"	
"The Semantic Web"	23

Données:

```
@prefix dc10: <http://purl.org/dc/elements/1.0/> .
@prefix dc11: <http://purl.org/dc/elements/1.1/> .
_:a dc10:title "SPARQL Query Language Tutorial" .
_:a dc10:creator "Alice" .
_:b dc11:title "SPARQL Protocol Tutorial" .
_:b dc11:creator "Bob" .
_:c dc10:title "SPARQL" .
_:c dc11:title "SPARQL" .
```

Title "SPARQL Protocol Tutorial" "SPARQL" "SPARQL (updated)" "SPARQL Query Language Tutorial"

X	y
	"SPARQL (updated)"
	"SPARQL Protocol Tutorial"
"SPARQL"	
"SPARQL Query Language Tutorial"	

Séquences de solutions

- Les motifs de solution génèrent une séquence non ordonnée de solutions, chacun étant une fonction partielle de variables présentes dans les motifs vers des termes RDF.
- Ces solutions sont ensuite traitées comme une séquence sur laquelle on peut appliquer des opérateurs (modificateurs)
 - Order by : permet de trier les solutions
 - Projection : choix des variables (SELECT)
 - Distinct : éliminer les doublons parmi les solutions
 - Reduced : peut éliminer des solutions non uniques
 - Offset : indique la position où commencer dans la séquence de solutions
 - Limit : restreint le nombre de solution

ORDER BY

```
PREFIX foaf: <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/>
SELECT ?name
WHERE { ?x foaf:name ?name }
ORDER BY ?name
PREFIX foaf: <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/>
PREFIX : <a href="http://example.org/ns#">http://example.org/ns#></a>
SELECT ?name
WHERE { ?x foaf:name ?name ; :empId ?emp }
ORDER BY ?name DESC(?emp)
```

DISTINCT

```
Données:
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
_:a foaf:name "Alice" foaf:mbox <mailto:alice@work.example> .
_:x foaf:name "Alice" foaf:mbox <mailto:smith@work.example> .
Requête:
PREFIX foaf: <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/>
SELECT DISTINCT ?name
WHERE { ?x foaf:name ?name }
Résultat
                       name
                      "Alice"
```

OFFSET et LIMIT

- OFFSET n : commencer à la solution n+1. OFFSET 0 n'a pas d'effet.
- LIMIT n : limite à n le nombre de solutions
- S'utilisent avec la clause ORDER BY

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?name
WHERE { ?x foaf:name ?name }

ORDER BY ?name
LIMIT 5

OFFSET 10
```

Le résultat de cette requête aura au maximum 5 solutions, à partir de la 11ème dans la séquence de solutions.

Fonctions

- Logique : !, &&, ||
- Math: +, -, *, /
- Comparaison : =, !=, >, <, ...
- Tests (SPARQL): isURI, isBlank, isLiteral, bound
- Autres (SPARQL): str, lang, datatype
- sameTerm, langMatches, regex

Exemples de requêtes

Données:

```
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix : <http://example.org/book/> .
@prefix ns: <http://example.org/ns#> .
:book1 dc:title "SPARQL Tutorial" .
:book2 dc:title "The Semantic Web" .
:book2 ns:price 23 .
```

Quantification existentielle

- OPTIONAL + bound() permet d'exprimer la quantification existentielle
- Bound() renvoie true si la variable est liée, false sinon.

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX ns: <http://example.org/ns#>
SELECT ?title ?price
WHERE { ?x dc:title ?title .
         OPTIONAL { ?x ns:price ?price } .
         FILTER ( bound(?price)) }
```

title	price
"The Semantic Web"	23

Quantification existentielle

• OPTIONAL + ! bound() permet de tester qu'un motif de graphe n'est pas exprimé.

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX ns: <http://example.org/ns#>
SELECT ?title ?price
WHERE { ?x dc:title ?title .
         OPTIONAL { ?x ns:price ?price } .
         FILTER ( !bound(?price)) }
```

title	price
"SPARQL Tutorial"	

Négation

- 2 façons d'exprimer la négation
 - FILTER NOT EXIST { ...} teste la non-existence d'un motif
 - MINUS permet de retirer des solutions provenant d'un autre motif de graphe

FILTER NOT EXISTS

```
:a foaf:givenName "Alice" .
_:b foaf:givenName "Bob" .
:b dc:date "2015-05T05:05:05"^^xsd:dateTime .
PREFIX foaf: <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/>
PREFIX dc: <a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/>
SELECT ?name
WHERE { ?x foaf:givenName ?name .
  FILTER NOT EXISTS { ?x dc:date ?date }}
Résultat :
                   name
                  "Alice"
```

MINUS

```
:a foaf:givenName "Alice" .
_:b foaf:givenName "Bob" .
:b dc:date "2015-05T05:05:05"^^xsd:dateTime .
PREFIX foaf: <a href="http://xmlns.com/foaf/0.1/">
SELECT ?o
WHERE { ?s ?p ?o .
  MINUS { ?s foaf:givenName "Bob" .}
Résultat:
                 0
               "Alice"
```

Remarque : La deuxième expression du MINUS doit contenir au moins une variable. On retire de la première expression les résultats de la deuxième dans lesquels les variables 43 communes ont la même valeur.

FILTER NOT EXISTS vs MINUS

```
SELECT
{ ?s ?p ?o
      FILTER NOT EXISTS { ?x ?y ?z }
Ne renvoie jamais de solution
SELECT
{ ?s ?p ?o
     MINUS { ?x ?y ?z }
Ici, il n'y a pas de variable commune, le MINUS ne retire aucune solution.
```

ASK

• Renvoie un booléen indiquant si un motif d'interrogation correspond ou non.

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
_:a foaf:name "Alice" .
_:a foaf:homepage <http://work.example.org/alice/> .
:b foaf:name "Bob" .
_:b foaf:mbox <mailto:bob@work.example> .
Requêtes:
  PREFIX foaf: <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/>
  ASK { ?x foaf:name "Alice" }
renvoie true
  PREFIX foaf: <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/>
  ASK { ?x foaf:name "Alice" ;
  foaf:mbox <mailto:alice@work.example>}
renvoie false
```

CONTRUCT

• CONSTRUCT renvoie un graphe RDF décrit par un gabarit de graphe.

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
        foaf:name "Alice".
:a
    foaf:mbox <mailto:alice@example.org> .
:a
PREFIX foaf: <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/>
PREFIX vcard: <a href="http://www.w3.org/2001/vcard-rdf/3.0#">http://www.w3.org/2001/vcard-rdf/3.0#</a>
CONSTRUCT {<http://example.org/person#Alice> vcard:FN ?name}
              { ?x foaf:name ?name }
WHERE
Crée des propriétés vcard à partir des informations FOAF:
@prefix vcard: <http://www.w3.org/2001/vcard-rdf/3.0#> .
<http://example.org/person#Alice> vcard:FN "Alice" .
```

CONSTRUCT

• Il est possible de créer un graphe avec des nœuds anonymes

CONSTRUCT

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX vcard: <a href="http://www.w3.org/2001/vcard-rdf/3.0#">http://www.w3.org/2001/vcard-rdf/3.0#</a>
CONSTRUCT { ?x vcard:N _:v .
             :v vcard:givenName ?gname .
             :v vcard:familyName ?fname }
WHERE {
{?x foaf:firstname ?gname } UNION {?x foaf:givenname
  ?gname } .
{?x foaf:surname ?fname } UNION {?x foaf:family_name
  ?fname } .
```

CONSTRUCT

• Résultat:

```
@prefix vcard: <http://www.w3.org/2001/vcard-rdf/3.0#> .
:v1 vcard:N
                      :x .
:x vcard:givenName "Alice" .
_:x vcard:familyName "Hacker" .
:v2 vcard:N
                      :Z.
:z vcard:givenName
                      "Bob" .
:z vcard:familyName "Hacker" .
                                            Hacker
                       vcard:familyName
          vcard:N
                                            Bob
                             vcard:givenName
                            vcard:familyName Hacker
          vcard:N
                                                         49
                                            Alice
                           vcard:givenName
```

Conclusion

- Interrogation de graphes sémantiques
 - Définition de motifs
 - Recherche de données correspondant au motif dans le graphe (appariement de graphes)
- Autres fonctionnalités
 - Règles d'inférence : pour déduire des informations
 - Nouveaux opérateurs et fonctions de filtre (négation, agrégats..) (SPARQL1.1)
 - Requêtes imbriquées
 - Chemins de propriétés

– ...

```
@prefix : <a href="http://dbpedia.org/resource/">http://dbpedia.org/resource/>...
    @prefix rdf: <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
    @prefix rdfs: <a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#>.
                                                                              #movies
    @prefix foaf: <a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/">.
                                                                               :drive
                                                                                 a :movie;
#actors
                                           #plays
                                                                                 :budget "12";
                                           :ryan_gosling
:ryan_gosling
                                                                                 :year "2012";
                                              :plays :drive,
  a foaf:actor;
                                                                                 :rating "7.8";
                                           :gangster_squad.
  rdfs:label "Ryan Gosling";
                                                                                 :seen "352000".
  :birthDate "1980-11-12";
                                           :Christina Hendricks
   :countryOfBirth "Canada".
                                                                              :madmen
                                              :plays :drive,
                                                                                 a :movie;
                                           :madmen.
:Christina_Hendricks
                                                                                 :budget "8";
  a foaf:actor;
                                                                                 :year "2007";
  rdfs:label "Christina Hendricks"; :Christina_Hendricks
                                                                                 :rating "7.2";
  :birthDate "1975-05-03".
                                              :plays
                                                                                 :seen "96000".
                                           :gangster_squad,
:Carey_Mulligan
                                           :never_let.
                                                                              :gangster_squad
  a foaf:actor;
                                                                                 a:movie;
  rdfs:label "Carey_Mulligan";
                                                                                 :budget "15";
  :birthDate "1985-05-28";
                                                                                 :year "2013";
  :countryOfBirth "UK".
                                                                                 :rating "6.8";
                                                                                 :seen "160000".
```

Exemple 1

select ?actor ?movie where{ ?actor :plays ?movie. ?movie :budget "12"}

Exemple 2

```
select ?actor ?birthCount
where
{ ?actor :plays :drive .optional { ?actor :countryOfBirth ?birthCount } }
```

Exemple 3

Les acteurs qui ont joué dans le même film.

Le résultat de la requête ne doit pas contenir de doublon :

actor1	actor2
:Christina_Hendricks	:ryan_gosling