points update :

1. après executer le méthode de cherche du ?ps, si le ?ps est <<http://dbpedia.org/ontology/wikiPageWikiLink>> , je fait le requête

 SELECT ?p (COUNT(?p) AS ?pTotal)

 where{

 ?s <http://dbpedia.org/ontology/wikiPageWikiLink> os

 ?s ?p os .

 }

 GROUP BY ?p

 ORDER BY DESC(?pTotal)

 pour trouver le prédicat substitut de <<http://dbpedia.org/ontology/wikiPageWikiLink>>.

 par example: pour le tripet:

 sc: <http://fr.dbpedia.org/resource/Catégorie:Lanterne\_des\_morts>

 ps: <http://dbpedia.org/ontology/wikiPageWikiLink>

 cs: <[http://fr.dbpedia.org/resource/Monument\_historique\_(France)](http://fr.dbpedia.org/resource/Monument_historique_%28France%29)>

 apres le requête

 SELECT ?p (COUNT(?p) AS ?pTotal)

 where{

 ?s <http://dbpedia.org/ontology/wikiPageWikiLink> <http://fr.dbpedia.org/resource/ Monument\_historique\_(France)> .

 ?s ?p <http://fr.dbpedia.org/resource/Monument\_historique\_(France)> .

 }

 GROUP BY ?p

 ORDER BY DESC(?pTotal)

 le résultat est

|  |  |
| --- | --- |
| **p** | **pTotal** |
| http://dbpedia.org/ontology/wikiPageWikiLink | 21826 |
| http://dbpedia.org/ontology/heritageRegister | 298 |
| http://fr.dbpedia.org/property/classement | 174 |
| http://fr.dbpedia.org/property/protection | 4 |
| http://dbpedia.org/ontology/wikiPageRedirects | 3 |
| http://fr.dbpedia.org/property/statut | 2 |
| http://dbpedia.org/ontology/knownFor | 2 |
| http://dbpedia.org/ontology/activity | 2 |
| http://fr.dbpedia.org/property/autreProtection | 1 |
| http://dbpedia.org/ontology/product | 1 |
| http://dbpedia.org/ontology/created | 1 |

alors je prends le prédicat <<http://dbpedia.org/ontology/heritageRegister>> comme un substitut de <http://dbpedia.org/ontology/wikiPageWikiLink> , et le triplet devient:

 sc: <http://fr.dbpedia.org/resource/Catégorie:Lanterne\_des\_morts>

 ps: <http://dbpedia.org/ontology/heritageRegister>

 cs: <http://fr.dbpedia.org/resource/Monument\_historique\_(France)>

normalement je prends le ?p le plus utilise sauf les deux: <http://dbpedia.org/ontology/wikiPageWikiLink> et <<http://dbpedia.org/ontology/wikiPageRedirects>>

2. J’ai modifié le requête pour trouver le ?os,

Pour le sous-catégorie ?sc donnée, je calcule les ?o par

?s ?p ?sc .

?s2 <<http://dbpedia.org/ontology/wikiPageWikiLink>> ?o

le méthode ?s <http://[dbpedia.org/ontology/type](http://dbpedia.org/ontology/type)> ?sc . ne s’applique qu’au Catégorie: Édifice-type.

Et j’ajoute aussi un seuil au nombre total de ?o.

3. Dans le programme du jena, j’ajoute le requête pour detecter si un triple deja existe dans Dbpedia:

ask{ sc ps os }

Les résultats affichent dans un tableau avec un nouveau valeur exist(true/false),

prend leCatégorie:Événement\_sportif\_international comme un exemple

|  |  |  |  |
| --- | --- | --- | --- |
| sc | ps | os | exist |
| <http://fr.dbpedia.org/resource/Dodécaphonisme> | <http://dbpedia.org/ontology/occupation> | <http://fr.dbpedia.org/resource/Compositeur> | 0 |
| <http://fr.dbpedia.org/resource/Playback> | <http://dbpedia.org/ontology/nationality> | <http://fr.dbpedia.org/resource/États-Unis> | 0 |
| <http://fr.dbpedia.org/resource/Featuring>  |  <http://fr.dbpedia.org/property/format> | <http://fr.dbpedia.org/resource/Album\_studio> | 0 |
| <http://fr.dbpedia.org/resource/Mélisme> |  <http://dbpedia.org/ontology/author>  | <http://fr.dbpedia.org/resource/Johann\_Sebastian\_Bach>  | 0 |
| <http://fr.dbpedia.org/resource/Suspension\_acoustique> | <http://fr.dbpedia.org/property/langue> | <http://fr.dbpedia.org/resource/Latin> | 0 |
| <http://fr.dbpedia.org/resource/Jeu\_de\_scène>  |  <http://dbpedia.org/ontology/nationality>  | <http://fr.dbpedia.org/resource/États-Unis>  | 0 |
| <http://fr.dbpedia.org/resource/Tyrolienne\_(chant)>  | <http://dbpedia.org/ontology/land> |  <http://fr.dbpedia.org/resource/Tyrol>  | 0 |

Apres que je teste cet algorithme sur différents catégories, la plupart des triples obtenues n’existent pas dans Dbpedia.

En meme temps, les résultats seront aussi sauvegardé dans des fiches en format json. Pour chaque catégorie, il cree deux fichiers, en séparant les résultats qui existent et les autres qui n’existent pas.

Les noms du fichier sont crées selon les different catégories.

par example:

Category\_Région\_d’Europe\_existRes.json

Category\_Région\_d’Europe\_notExistRes.json

Category\_Événement\_sportif\_international\_existRes.json

Category\_Événement\_sportif\_international\_notExistRes.json