

Projet “Impact scientifique”

Pour ce projet vous pouvez travailler en groupes de 2 ou 3 personnes.

Vous pouvez écrire votre code en Jena ou RDFlib, le choix du langage de programmation est libre

Pour le rendu vous devriez écrire un petit rapport avec la réponse à chaque question du projet et rendre celui-ci avec le code (le tout en un seul fichier .zip à soumettre sur Moodle)

Date limite pour le rendu: **15 avril 2022** à minuit sur **moodle**:
<https://moodlelms.univ-paris13.fr/course/view.php?id=1081>

Description:

Dans DBPedia nous avons des informations structurées autour de personnages du sport, culture, politique, science, etc.

Observons par exemple la fiche DBPedia de **Marie Curie**:
https://dbpedia.org/page/Marie_Curie

Vous pouvez trouver des informations comme la/les discipline/es dans laquelle le scientifique a travaillé:

[dbo:academicDiscipline](#)

- [dbr:Chemistry](#)
- [dbr:Physics](#)

Savoir s'il s'agit d'un scientifique grâce à la propriété **rdf:type** si l'objet est **dbo:Scientist**

Savoir si le scientifique est une femme via la relation **rdf:type** avec pour objet **yago:Woman110787470**

Vous avez en plus des relations qui permettent de savoir quels scientifiques ont été influencés par celui de la fiche et vice versa:

- **dbo:influencedBy** <https://dbpedia.org/ontology/influencedBy> *
- **dbo:influenced** <https://dbpedia.org/ontology/influenced>
- **dbp:influences** <https://dbpedia.org/property/influences> *
- **dbp:influenced** <https://dbpedia.org/property/influenced>

Objectifs:

Q.1) Ecrire une fonction qui permet de chercher tous les scientifiques qui ont été influencés par quelque autre scientifique (en paramètre). Ce point ne se limite pas aux influences directes mais aussi celles **indirectes**. Faites attention au risque de références **circulaires** (ce qui risque de produire des boucles infinies).

Q.2) Ecrire une fonction qui, étant données les noms de deux scientifiques, permet de savoir s'il existe une connexion entre eux et à que distance (idée similaire à celui de la [distance de Kevin Bacon](#))

Q.3) Ecrire une fonction qui permet d'obtenir une liste de scientifiques avec le nombre de domaines différentes qui ont été impactés par leur recherches (par exemple, dans le cas de Marie Curie, au moins la physique et la chimie, mais aussi des autres si entre les scientifiques influencés par Marie Curie il y en a certains qui ont travaillé dans des disciplines différentes).

Q.4) Ecrire une fonction qui permet d'obtenir une liste de femmes scientifiques les plus influentes (c'est à dire, qui ont influencés le plus de monde et le plus de disciplines différentes) de l'histoire.

Suggestions:

Bien sûr vous pourriez télécharger les données de DBPedia, par contre on vous conseille d'utiliser le point d'accès SPARQL: <http://dbpedia.org/sparql>

Nous avons vu comment inclure des requêtes SPARQL dans Jena ou RDFlib dans les derniers TPs