

Rencontre Emvista

Présentation d'un article scientifique

Rodrique Kafando

Özcan, F., Lei, C., Quamar, A., Efthymiou, V. (2021, June). Semantic enrichment of data for AI applications. *In Proceedings of the Fifth Workshop on Data Management for End-To-End Machine Learning* (pp. 1-7).

07.01.2022



Contexte

Table

Rank	Name	No. of reigns	Combined days
1	Lou Thesz	3	3,749
2	Ric Flair	8	3,103
3	Harley Race	7	1,799
4	Dory Funk Jr.	1	1,563
5	Dan Severn	2	1,559
6	Gene Kiniski	1	1,131

Example questions

#	Question	Answer
1	Which wrestler had the most number of reigns?	Ric Flair
2	Average time as champion for top 2 wrestlers?	AVG(3749,3103)=3426
3	How many world champions are there with only one reign?	COUNT(Dory Funk Jr., Gene Kiniski)=2
4	What is the number of reigns for Harley Race?	7



Questions/Answers application

Enrichir sémantiquement des données structurées à partir de graphes de connaissances inter-domaines et des ontologies de domaine, afin de mieux interpréter les requêtes et fournir des réponses plus complètes et désambiguïsées.

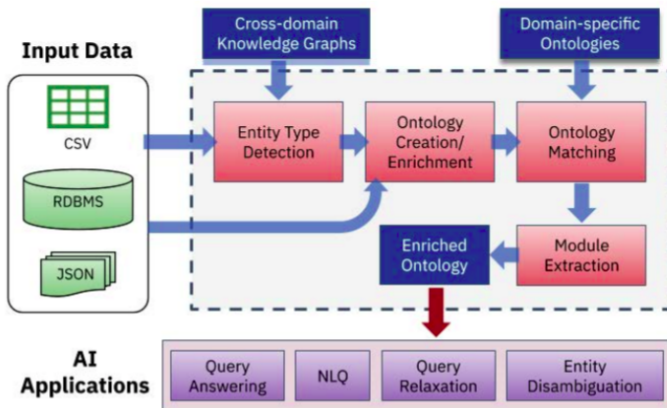
Problème posé

Comment enrichir une base de données quelconque avec des ressources sémantiques (graphes de connaissances, ontologies de domaines, etc.) externes ?

- Le principe repose sur le fait que dans une base de données (ex: SGBDR), il est possible d'identifier différents types d'entités relatives à un domaine identifié ou non.
 - ▶ en générant une ontologie spécifique à cette base de données ;
 - ▶ en trouvant les correspondances entre cette ontologie avec des ontologies génériques (associé à différents types d'entités) ;
 - ▶ et en enrichissant les différents niveaux de hiérarchies avec des calculs de similarité.

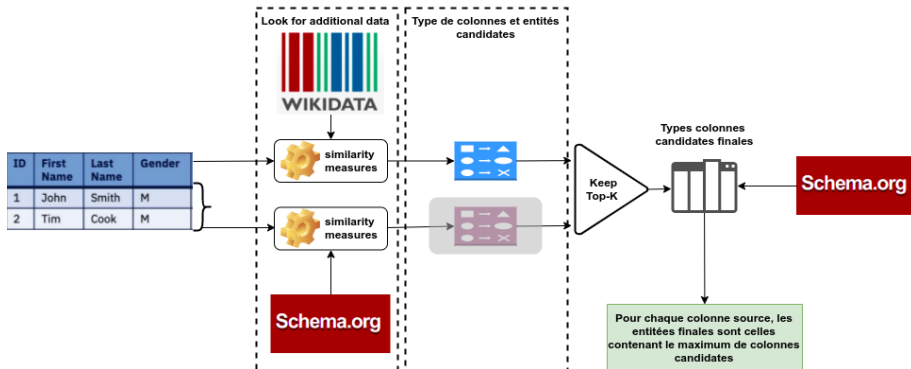
Approche méthodologique

- 4 principales étapes :



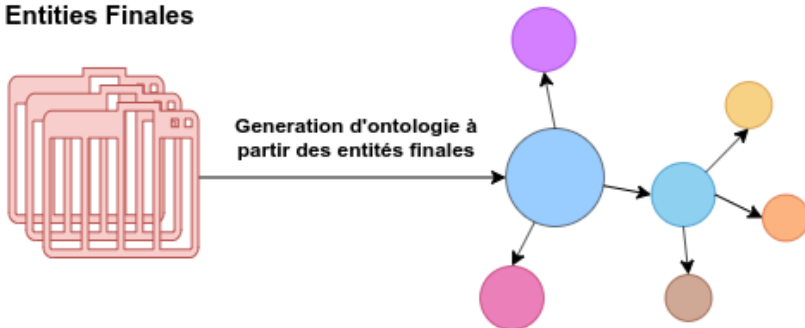
Approche méthodologique

Step 1 : Entity Type Detection



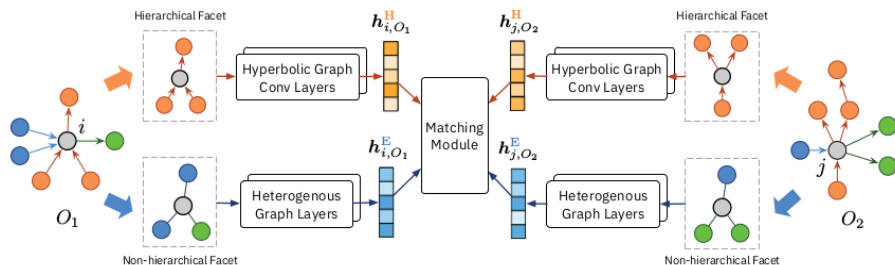
Step 2 : Ontology Creation/Enrichment

Entités Finales



Ontologie construite en se basant sur : les entités avec leurs propriétés et les relations existant entre elles

Step 3 : Ontology to Ontology Matching



Step 4 : MODULAR ONTOLOGY REUSE

- extraire un module minimal O'_2 dans O_2 : Ne conserver que le plus petit sous-ensemble possible de O_2 qui est pertinent vis-à-vis de O_1 , sans perdre la sémantique;
- fusionner O_1 et O'_2

Dans le but de mieux interpréter les requêtes et fournir des réponses plus complètes et désambiguïsées dans certaines applications de l'IA, les auteurs ont proposé dans cet article, une approche permettant de:

- construire ou de générer des ontologies à partir de données structurées en utilisant des ressources externes (bases de connaissances);
- enrichir cette ontologie en utilisant une ontologie d'un domaine spécifique

Merci pour votre attention!