

Rapport du projet « Ontologie sur l'invertissement immobilier dans les villes chinoises »

Web sémantique et big-data

Shilin XIE, Yangduo ZOU, Yuyan QIAN
21102552, 21209518, 21108372

Table des matières

1	Introduction	2
2	Définition et état de l'art	2
3	Méthodologie	3
3.1	Création des classes, des sous-classes et des instances	3
3.2	Création des propriétés : object property, data property	4
3.3	Les inférences et la consistance	4
4	Requête SPARQL	5
5	Conclusion	6

1 Introduction

L'arrivée des ontologies dans le domaine informatique, nous aide à résoudre des problèmes rencontrés lors de la représentation et la manipulation des connaissances dans les systèmes informatiques. Dans le cadre de Web sémantique, nous avons exploré les possibilités de l'ontologie ainsi que le logiciel protégé. Afin de pratiquer les savoirs, nous avons repris le sujet du *Semantic MediaWiki* : les villes chinoises. Un échantillon de villes est choisi dans le but de mieux présenter la diversité urbaine en Chine. Nous verrons dans ce rapport le sujet détaillé du projet et les étapes établies, en commençant par une simple comparaison avec les autres ontologies de villes. Ensuite, nous expliquerons la construction de notre projet en donnant des exemples, suivie par la mise en pratique SPARQL. Pour conclure, nous analyserons l'accomplissement de ce projet.

2 Définition et état de l'art

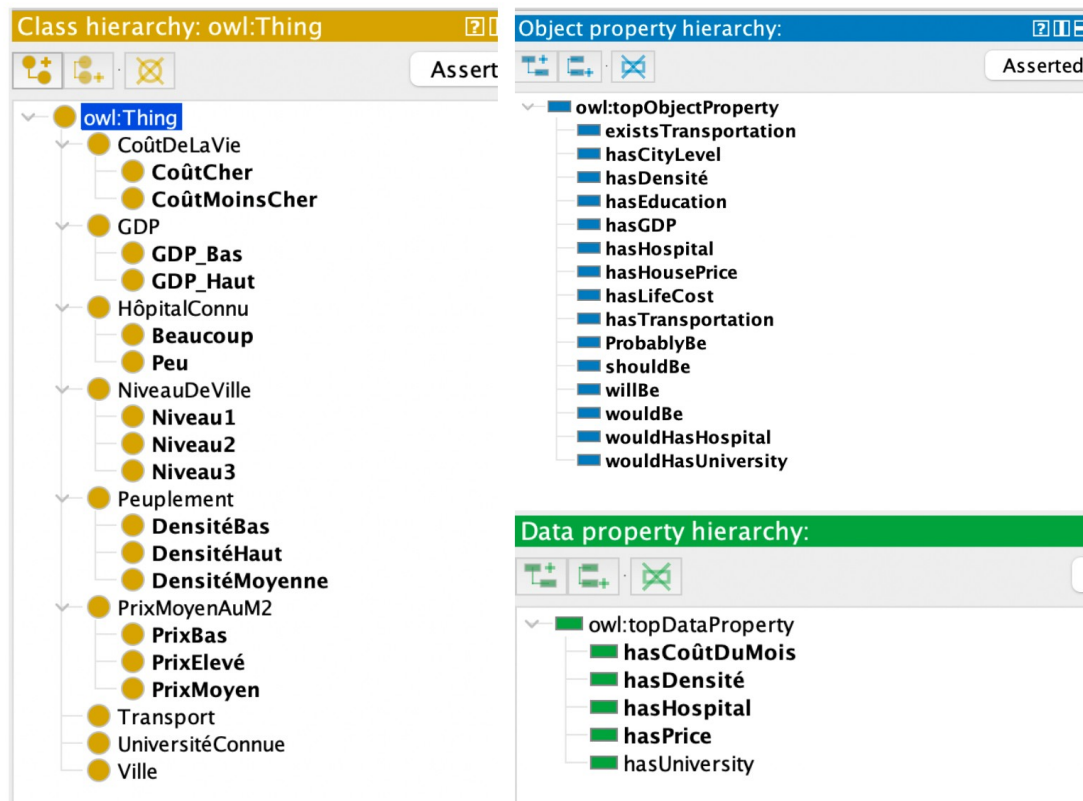
Comme ce que nous avons indiqué dans la description du sujet, « Ville » est un sujet trop général pour aborder. Autour de ce concept, nous avons planifié d'explorer les problèmes urbains sous différents niveaux. Pour mettre en place ce projet, nous nous avons concentré sur l'investissement immobilier dans les villes chinoises. Étant donné la complexité, nous supposons que l'investisseur prend en compte de nombreux facteurs, tels que le coût de la vie, le produit intérieur brut (GDP en anglais), etc. Dans ce contexte, l'utilisation d'une ontologie organisée permettra de faciliter la prise de décision en offrant une représentation sémantique des connaissances relatives aux villes.

Au cours de l'élaboration thématique, nous nous sommes référés à un projet existant *Sustainable Cities and Human Settlements, v1.0*¹. La création d'une ontologie pour l'investissement immobilier dans les villes chinoises est plus spécifique au domaine de l'immobilier. Cependant, les deux ontologies peuvent être liées car les éléments clés sont partagés, tels que le transport, la population, etc. À partir de ces parties communes, nous voudrions le développer en détaillant les villes et en multipliant les relations entre les nœuds en vue de maximiser la fonction de raisonnement.

En résumé, l'utilisation de Protégé pour la création d'ontologies pour l'investissement immobilier dans les villes chinoises peut être utile pour les investisseurs, les développeurs immobiliers et les décideurs politiques.

1. https://github.com/IemProg/IA301_SymbolicAI_Ontologies

3 Méthodologie



3.1 Création des classes, des sous-classes et des instances

En étape initiale, nous avons créé un glossaire des termes concernant ce sujet. Tout les éléments sont transformés en classe et sous-classe correspondante comme ci-dessus (9 classes, 15 sous-classes, 30 instances). Pour notre thème, de nombreux éléments ne peuvent être instanciés directement. Par exemple, les prix moyens au mètre carré de l'appartement pour chaque ville sont des chiffres spécifiques et, dans ce cas, nous regroupons ces chiffres spécifiques en différents intervalles (Prix-bas, prix-moyen et prix-élevé) que nous définissons comme des sous classes. *Prix-bas* a une instance *prix 30000-* et ce prix correspond à la ville de niveau III (*Prix30000- should be Niveau3Ville*). En même temps, une data propriété *hasPrice* est attribuée à cette instance avec une valeur de type int 26000 (*Prix30000- hasPrice 26000*). En créant un pont d'intervalle entre les chiffres et les classes de facteurs, nous avons essayé de catégoriser et d'individualiser les classes. Cette méthode nous a permis d'éviter l'augmentation des nombres d'entités, lorsque le nombre de villes augmente et que varient le PIB, les prix de l'immobilier et les autres facteurs correspondants.

Ce problème d’instanciation est dû à la numérisation excessive parue dans notre sujet. Et dans la vie réelle, il est également difficile de trouver des moyens adéquats pour instancier de tels problèmes, de sorte que cette dé-numérisation peut également s’appliquer à d’autres pratiques ontologiques.

3.2 Création des propriétés : object property, data property

La data propriété crée la relation entre l’entité et ses propres données, et elle nécessite la définition du domaine et de le range, qui est le type des données. Par rapport à celle-ci, la objet propriété est consacrée à relier deux classes. Dans notre ontologie, *hasLifeCost* a un sujet de ville et un objet de coût de la vie. Pour affiner cette relation, une restriction est ajoutée dans la partie de range : au plus un seul coût de la vie. Pour clarifier la data propriété, nous voudrais citer une définition de notre projet, soit le triple de *hasPrice*, le prix moyen au mètre carré a un prix de type integer.

Dans cette partie, nous n’avons pas beaucoup décrit la caractéristique de l’objet propriété, telle que transitive, symétrique, puisque la relation a causé des problèmes de raisonnement dans la complexité des relations métropolitaines. En raison de cela, nous avons choisi de privilégier la cohérence globale.

3.3 Les inférences et la consistance

Explanation 1 ☐ Display laconic explanation

Explanation for: Shanghai Type DensitéHaut

- Shanghai hasHousePrice Prix50000+
- Prix50000+ Type PrixElevé
- DensitéHaut EquivalentTo hasHousePrice some PrixElevé

La principale fonction de ce logiciel protégé est de représenter les relations invisibles par un raisonnement logique. Une ontologie comportant des erreurs internes ne sera pas correctement déduite. Par exemple, l’instance Shanghai a un prix moyen au mètre carré de plus de 50000 RMB et l’intervalle de plus de 50000 RMB correspond à un niveau de prix élevé. Avec la dernière attestation, une forte densité de population aurait pu être un prix moyen au mètre carré élevé. Nous pouvons en déduire que Shanghai est une ville à forte densité d’importations humaines.

4 Requête SPARQL

1. Recherche la coût de la vie à chaque ville

SPARQL query:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX base: <http://www.semanticweb.org/shailynxie/ontologies/2023/2/untitled-ontology-6#>

SELECT ?city ?lifecost
WHERE { ?city base:hasLifeCost ?lifecost }
```

city	lifecost
Kunming	Coût5000-
Shanghai	Coût10000+
Chengdu	Coût5000-
Guangzhou	Coût10000+
Sanya	Coût5000-
Beijing	Coût10000+

Execute

4. Recherche les villes où se trouvent les universités et leurs nombres

SPARQL query:

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX base: <http://www.semanticweb.org/shailynxie/ontologies/2023/2/untitled-ontology-6#>

SELECT ?city ?Nombre ?nb
WHERE { ?city base:hasEducation ?Nombre.
?Nombre base:hasUniversity ?nb
FILTER(?nb > 0) }
```

city	Nombre	nb
Guangzhou	Nombre2	"2"^^<http://www.w3.org/2001/XMLSchema#integer>
Chengdu	Nombre2	"2"^^<http://www.w3.org/2001/XMLSchema#integer>
Beijing	Nombre8	"8"^^<http://www.w3.org/2001/XMLSchema#integer>
Shanghai	Nombre4	"4"^^<http://www.w3.org/2001/XMLSchema#integer>

Execute

2. Recherche le niveau de chaque ville et le trie en ordre croissant

SPARQL query:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX base: <http://www.semanticweb.org/shailynxie/ontologies/2023/2/untitled-ontology-6#>

SELECT ?city ?level
WHERE { ?city base:hasCityLevel ?level }
ORDER BY ?level
```

city	level
Shanghai	Niveau1Ville
Beijing	Niveau1Ville
Guangzhou	Niveau1Ville
Chengdu	Niveau2Ville
Kunming	Niveau2Ville
Sanya	Niveau3Ville

Execute

5. Recherche le prix moyen au mètre carré pour chaque ville et leurs niveaux

SPARQL query:

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX base: <http://www.semanticweb.org/shailynxie/ontologies/2023/2/untitled-ontology-6#>

SELECT ?city ?HousePrice ?prix ?level
WHERE { ?city base:hasCityLevel ?level.
?city base:hasHousePrice ?HousePrice.
?HousePrice base:hasPrice ?prix }
```

city	HousePrice	prix	level
Beijing	Prix50000+	"63000"^^<http://www.w3.org/2001/XMLSchema#float>	Niveau1Ville
Shanghai	Prix50000+	"63000"^^<http://www.w3.org/2001/XMLSchema#float>	Niveau1Ville
Guangzhou	Prix50000+	"63000"^^<http://www.w3.org/2001/XMLSchema#float>	Niveau1Ville
Chengdu	Prix30000-50000	"38000"^^<http://www.w3.org/2001/XMLSchema#float>	Niveau2Ville
Kunming	Prix30000-	"26000"^^<http://www.w3.org/2001/XMLSchema#float>	Niveau2Ville
Sanya	Prix30000-50000	"38000"^^<http://www.w3.org/2001/XMLSchema#float>	Niveau3Ville

Execute

3. Recherche les types de transport commun à Guangzhou

SPARQL query:

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX base: <http://www.semanticweb.org/shailynxie/ontologies/2023/2/untitled-ontology-6#>

ASK {
base:Guangzhou
base:hasTransportation base:Bateau
}
```

Result
True

Execute

6. Comparer le nombre des universités à Beijing et à Guangzhou

SPARQL query:

```
ASK {
{
base:Guangzhou base:hasUniversity ?Nombre_GZ.
base:Beijing base:hasUniversity ?Nombre_BJ.
FILTER(?Nombre_GZ > ?Nombre_BJ)
}
}
```

Result
False

Execute

5 Conclusion

Il est clair que notre ontologie contient encore des lacunes, et comme nous l'avons analysé précédemment, nous avons essayé d'adapter notre sujet à la construction d'ontologies, dans lequel nous avons également constaté qu'il est plus facile de construire des ontologies avec un sujet qui peut être instanciés. Cependant cela ne signifie pas que d'autres types de sujet ne peuvent pas construire des ontologies, et la méthode que nous utilisons résout ce problème de manière assez satisfaisante.

Dans ce rapport, nous avons fait un rapide tour des ontologies et une analyse courte de nos méthodes. Ce travail consistait en la construction d'une ontologie portant sur l'investissement immobilier dans les villes chinoises. Et ce projet nous a permis d'en apprendre davantage sur le fonctionnement des ontologies et sur les possibilités qu'elles offrent. Vu l'ontologie existante de ce thème, nous croyons que dans la domaine d'investissement, un système basé sur une telle ontologie va nous renseigner les conseils avec rapidité et précision sans avoir une expérience de travail dans cette domaine. Il suffirait de quelques informations seulement pour trouver le ville idéale à choisir.