

Rapport du projet Web sémantique et big-data

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

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 cette description, nous allons détailler le sujet de notre projet et les raisons du choix, en commençant par la présentation des capacités d'OWL. Ensuite, nous expliquerons la modélisation et l'utilisation de notre ontologie en donnant des exemples.

L'ontologie est d'origine la philosophie depuis le XIXème siècle, qui est définie comme l'étude des propriétés générales de ce qui existe. Les ontologies sont nées du besoin de représenter les connaissances dans les systèmes informatiques, de ce fait, elles sont toujours définies par rapport au processus général de la représentation des connaissances. En résumé, une ontologie est une spécification explicite d'une conceptualisation. Il repose sur des définitions précises de concepts d'un domaine et de leurs relations. Ces définitions sont essentiellement acceptables et partageables par les utilisateurs et interprétables par des machines. Les constituantes de base de conception d'ontologie incluent : des concepts, des classes et les propriétés de chaque instance définissant des caractéristiques diverses. OWL est un langage d'ontologie de réseau qui décrit les différentes entités ou concepts existant dans le monde réel, ainsi que les relations qui existent entre eux. Chaque entité est identifiée par un identifiant unique global, tout comme chaque personne a un numéro d'identification ; le deuxième consiste à utiliser des paires attribut-valeur pour décrire les caractéristiques internes de l'entité, à utiliser des relations pour connecter deux entités et décrire leur relation.

En passant à notre ontologie, « la ville » est un sujet très général et important. Autour de ce concept, nous pouvons explorer de nombreux problèmes urbains sous différents niveaux. Pour mettre en place ce projet, nous avons l'intention de construire une ontologie sur les villes chinoises. En plus, avec l'essor de l'économie, de plus en plus de prévoyants prennent part à l'investissement dans les domaines variées. La plupart d'entre eux partagent une préoccupation d'investir l'immobilier, vu que le besoin de logement est toujours indispensable pour l'homme du monde entier, particulièrement pour les habitants métropolitains. Cependant, l'investissement immobilier est une tâche tellement

complexe et compliquée. C'est la raison pour laquelle notre projet s'articulera autour de l'investissement immobilier dans différentes villes de Chine, plus précisément, nous nous concentrerons sur l'investissement d'appartements dans les villes. Étant donné la complexité, nous supposons que l'investisseur prend en compte de nombreux facteurs, tels que le coût de la vie, l'infrastructure, et l'environnement naturel 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.

Notre projet sera axé sur l'investissement d'appartements dans les différentes villes chinoises. Et l'ontologie sera créée dans le but d'aider les acheteurs potentiels à trouver facilement les villes où les appartements répondent le mieux à leurs besoins. Nous allons utiliser différentes sources pour créer cette ontologie, notamment des sites immobiliers, des données gouvernementales sur les villes et des rapports analytiques concernant le développement des villes chinoises. Le concept de ville englobe de nombreux éléments interconnectés. Ces éléments ne sont pas indépendants les uns des autres, mais plutôt en interaction constante, contribuant ainsi à façonner la ville et à la distinguer d'autres environnements urbains. La relation entre chaque élément et la ville est donc unique et complexe, rendant ainsi l'étude et la compréhension de ce concept fascinantes. En pratique, tous les éléments, nous les définissons comme des classes. Et ils influencent la décision des investisseurs. Il est préférable d'utiliser des triplets pour décrire ces données. L'ontologie permettra de filtrer les villes en fonction de critères de recherche tels que le climat, le prix moyen au mètre carré et la proximité des services publics, tout en fournissant des informations sur les différentes villes.

Les compétences de notre ontologie se basent sur les questions ci-dessus : Quels sont les critères importants pour les acheteurs d'appartements dans différentes villes ? Comment classer les villes selon différents critères tels que les espaces verts, les transports en commun, l'éducation supérieure, etc. ? Comment utiliser OWL pour représenter les informations sur les appartements et les villes de manière efficace ? Quelles sont les relations qu'on doit formaliser entre les facteurs et les villes afin de classer les lieux idéaux à investir ?

L'outil de construction « Protégé » est utilisé pour créer une bibliothèque d'ontologies. Cette bibliothèque est composée de deux parties principales. La première partie concerne les classes, chacune d'entre elles regroupant des instances qui appartiennent à cette catégorie. La seconde comprend deux types de propriétés : les propriétés d'objet, qui sont des attributs relationnels décrivant les relations entre deux entités, et les propriétés de données, qui sont des attributs intrinsèques décrivant les caractéristiques de l'entité elle-même.

Une fois que toutes les classes, relations et instances nécessaires auront été créées à l'aide du logiciel Protégé, les connexions entre elles deviendront clairement visibles, facilitant ainsi les requêtes de l'utilisateur. De plus, la fonction de raisonnement peut être

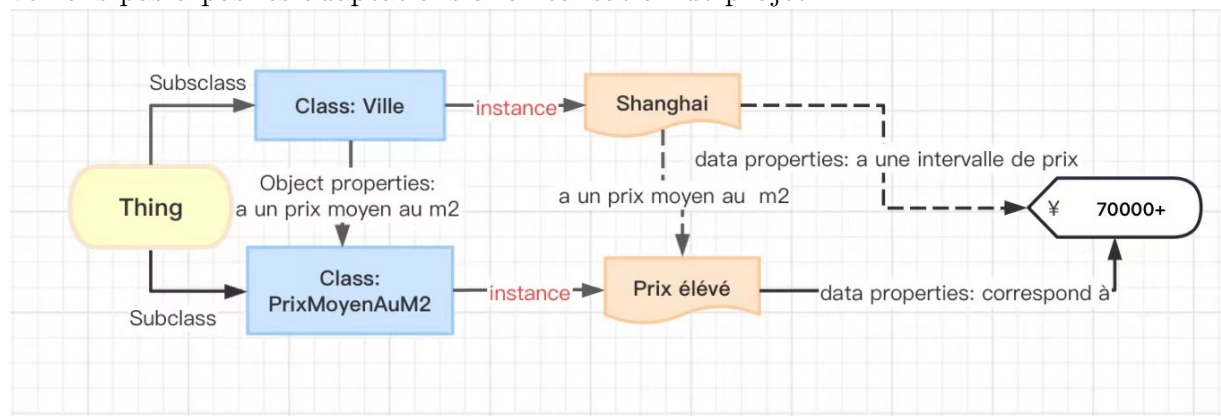
utilisée pour permettre à l'ontologie de déduire les villes les plus appropriées en fonction des besoins d'investissement individuels, puis de les classer en conséquence, ce qui aide l'investisseur à prendre des décisions plus éclairées et efficaces.

Voici une démonstration simple des classes dans notre projet :

Classe	Description de classe	Exemples des instances
Ville	Les différentes villes en Chine	Beijing, Shanghai, GuangZhou, etc
PrixMoyenAuM2	Le niveau du prix moyen par mètre carré dans une ville	haut, moyen, bas
TauxDesPlusValues	Augmentation ou diminution de la valeur des appartements dans une ville par rapport à l'année dernière	haut, moyen, bas
CoûtDeLaVie	Le niveau du coût de la vie dans une ville	haut, moyen, bas
SystèmeMédical	Les conditions de système médical dans une ville	hôpital général, hôpital spécial
Tourisme	Le développement du tourisme dans une ville	
Education	Le développement de l'éducation dans une ville, par exemple, une ville existe des universités connues ou pas	
Administration	Les différentes politiques d'achat d'appartement dans une ville	
Commerce	Le développement du commerce dans une ville	
Catastrophe Naturelle	L'environnement naturel	séisme, inondation, etc
Climat	Les différents types de climat	climat tropical, climat océanique, etc
GDP	Les trois niveaux de GDP	haut, moyen, bas
Géographie	Les caractéristiques géographiques de la ville	
Peuplement	L'ensemble des populations des espèces appartenant souvent à un même groupe taxonomique	
Population	Un ensemble d'individus dans une ville	
Superficie	Étendue de sol de cette ville	

Transport	Les conditions de transport dans une ville	métro, aéroport, bus, etc
Verdissement	Le niveau des espaces verts dans une ville	

En réalité, les facteurs mentionnés ci-dessus sont généralement exprimées sous forme quantitative, par exemple, le prix moyen au mètre carré d'une appartement est une valeur numérique concrète qui ne peut pas être considérée comme une instance. Nous pourrions donc imaginer trois instances pour dénumérer cette classe «PrixMoyenAuM2» : « prix élevé », « prix moyen » et « prix bas ». La classe « ville » aura un *object properties* a « a une intervalle de prix » qui relierait cette classe à une autre classe, « PrixMoyenAuM2 ». Comme l'exemple présenté ci-dessous, il existe plusieurs autres facteurs à déchiffrer. Nous verrons pas à pas les adaptations à la réalisation de projet.



Le but de ce projet est de construire une base de connaissances dans les domaines verticaux et de fournir des informations et des idées utiles pour choisir les villes pour l'investissement immobilier. Il peut être utilisé pour les questions et réponses dans le domaine de la sélection des villes d'investissement immobilier, de la recherche intelligente et de la recommandation : selon les besoins des différents utilisateurs, entrez les conditions de recherche et laissez le logiciel lui recommander les villes les plus appropriées, améliorez l'efficacité de trouver la maison des utilisateurs et aider les investisseurs à mieux comprendre la valeur d'investissement et le potentiel des villes pour prendre des décisions d'investissement plus éclairées.