

TP ISS Web Sémantique

Introduction

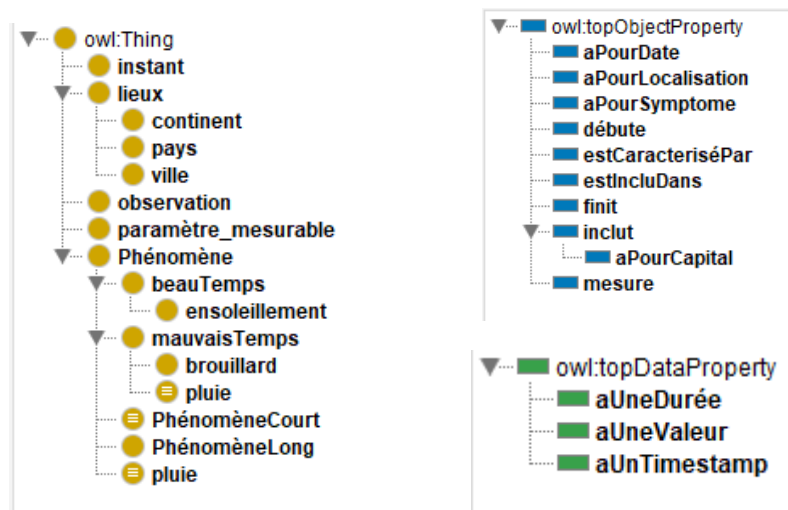
Actuellement, la majeure partie des pages internet sont programmées de manière à être lisibles pour l'Homme, lui permettant par la même occasion, de trouver, partager ou combiner des informations facilement. Cependant, ces tâches fastidieuses peuvent être réalisées par les machines, à condition de leur en donner les moyens ; c'est l'objectif du web sémantique. Le web sémantique permet aux machines de comprendre la sémantique, la signification de l'information sur le web. Via une structuration particulière des données, les machines peuvent ainsi effectuer des tâches automatisées (recherche, apprentissage...) mais également plus précises pour les utilisateurs.

Dans ce compte rendu de TP, nous allons voir comment fonctionne le web sémantique et notamment détailler le fonctionnement de son outil principal : l'ontologie.

1. Ontologie légère

Dans cette première partie, nous voulons développer une application de météorologie intelligente. Pour ce faire, nous décrivons une ontologie simple afin de fournir un vocabulaire permettant de décrire sémantiquement les données météorologiques.

Nous avons implémenté trois entités différentes pour cette ontologie : les classes (par exemple : beau temps, pluie, observations), les propriétés d'objets (exemple : inclus) et les propriétés de données (exemple : a une durée).



La différence entre les propriétés d'objet et de données vient du fait qu'une propriété d'objet relie deux classes, alors qu'une propriété de donnée relie une classe avec une donnée brute devant être typée. Par exemple, la propriété de donnée aUneValeur relie une observation à une valeur (dans notre cas de type integer), mais valeur n'est pas une classe.

Nous avons ensuite peuplé cette ontologie (voir fichier.owl).



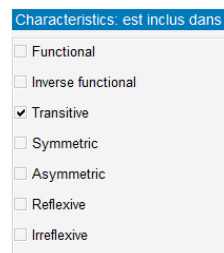
Remarque :

Lorsque nous avons créé Toulouse et France, nous ne les avons pas reliés à une classe. Cependant, lorsque nous avons rajouté la propriété estInclue entre Toulouse et la France, et que nous avons lancé le raisonneur, celui-ci a automatiquement associé aux individus Toulouse et France la classe lieu.

2. Ontologie lourde

En se basant sur l'ontologie légère que nous venons de créer, nous rajoutons des propriétés dans cette deuxième partie. Ces propriétés auront pour objectif d'enrichir nos données en déduisant des informations. Pour ce faire, nous ajoutons les propriétés suivantes :

- Toute instance de ville ne peut pas être un pays
- Un phénomène court est un phénomène dont la durée est de moins de 15 minutes
- Un phénomène long est un phénomène dont la durée est au moins de 15 minutes
- Un phénomène long ne peut pas être un phénomène court
- La propriété indiquant qu'un lieu est inclus dans un autre a pour propriété inverse la propriété indiquant qu'un lieu en inclut un autre.
- Si un lieu A est situé dans un lieu B et que ce lieu B est situé dans un lieu C, alors le lieu A est situé dans le lieu C



- A tout pays correspond une et une seule capitale

Characteristics: a pour capitale

| | |
|-------------------------------------|--------------------|
| <input checked="" type="checkbox"/> | Functional |
| <input type="checkbox"/> | Inverse functional |
| <input type="checkbox"/> | Transitive |
| <input type="checkbox"/> | Symmetric |
| <input type="checkbox"/> | Asymmetric |
| <input type="checkbox"/> | Reflexive |
| <input type="checkbox"/> | Irreflexive |

- Si un pays a pour capitale une ville, alors ce pays contient cette ville

Description: a pour capitale

Equivalent To

SubProperty Of

☒ inclus

- La Pluie est un Phénomène ayant pour symptôme une Observation de Pluviométrie dont la valeur est supérieure à 0.

Nous avons ensuite peuplé cette ontologie :

- La France est située en Europe

Rajout de l'individu Europe et de la propriété estInclueDans entre France et Europe

- Paris est la capitale de la France

Ajout de la propriété aPourCapitale entre Paris et France

- La Ville Lumière est la capitale de la France

Ajout de l'individu VilleLumière et de la propriété aPourCapital entre VilleLumière et France

- Singapour est une ville et un pays

Ajout de l'individu Singapour, en lui associant la classe ville ET pays.

Le raisonneur déduit alors que :

- A1 est un phénomène de type pluie (pluviométrie non nul).
- Paris est en Europe via la transitivité : le raisonneur « sait » que paris est la capitale de la France et que la France est en Europe. De plus, vu qu'un pays a une unique capitale, le raisonneur comprends que Paris est la ville lumière.

De plus, si Toulouse est déclarée comme la capitale de la France, le raisonneur déduit que Toulouse est la même ville que paris. En ce qui concerne Singapour, le raisonneur nous renvoie une erreur. En effet, la ville de Singapour et le pays étant deux entités différentes, le raisonneur ne prend pas en compte cette information fausse, assurant ainsi la cohérence des données. Pour décrire ce cas particulier, il faudrait déclarer un pays Singapour ainsi qu'une ville du même nom définie comme étant la capitale de ce pays.

3 Exploitation de l'ontologie

Par rapport aux données brutes issues du dataset, l'ontologie offre une structure bien définie permettant de relier les données entre elles. Cette hiérarchie permet également de s'assurer que les données peuplant l'ontologie sont cohérentes entre elles, par exemple à l'aide d'un raisonneur qui permet d'éviter les doublons, signaler les ambiguïtés...

Comme pour la manipulation via Protégé, l'API fait une distinction entre le label et l'URI : l'URI identifie de manière unique un individu qui peut avoir le même label qu'un autre.

On en déduit que l'API permet d'implémenter de manière rigoureuse une ontologie conçue en utilisant protégé. Ainsi, on peut imaginer une ontologie peuplée par le biais de requêtes HTTP et dont la structure est régulièrement testée via un raisonneur implémenté en Java.

En ce qui concerne le code, nous n'avons pas été en mesure d'implémenter l'intégralité de l'ontologie, nous avons seulement pu détailler les fonctions createPlace et createInstant permettant d'instancier un lieu dans notre ontologie et de le relier à son URI (idem pour les Instants, qui ont en plus une DataProperty associée).