



# Projet final

Web sémantique et Big data

---

## OBJECTIF

Ce travail pratique vous permettra de construire une ontologie sur le sujet de votre choix, de la peupler et d'exécuter des requêtes. Il faut que cette ontologie soit construite dans la perspective du Web sémantique dans le sens où il faut illustrer comment elle pourrait être utilisée, par exemple pour organiser un ensemble de données structurées en RDF et/ou annoter des ressources Web d'un domaine particulier et retrouver des informations complexes pour des applications.

## DESCRIPTION

Le travail est composé de trois étapes décrites à ci-dessous :

### Etape 1 : Phase initiale

1. **Formation d'une équipe de travail** : Constituer un groupe de travail de 3 personnes.
2. **Choix et description du sujet** : Choisir un sujet et le décrire (environ une à deux pages). Il faut que le sujet permette d'utiliser les différentes capacités d'OWL. Vous indiquerez aussi les raisons de votre choix, quels sont vos points de départ, expériences personnelles, documents textuels pour construire votre ontologie.
3. **Définition du domaine et sa portée** : De plus, et c'est essentiel, expliquer la tâche pour laquelle cette ontologie pourrait être construite en donnant des exemples de son utilisation. Quel domaine l'ontologie va-t-elle couvrir ? Dans quel but utiliserez-vous l'ontologie ?
4. **Questions de compétence** : Lister un certain nombre de questions de compétence.
5. **Rédaction du rapport** : Rédiger un rapport en indiquant et motivant le sujet de votre choix.

### Etape 2 : Phase de réalisation

1. **Etat de l'art** : Dans le rapport que vous rendrez, vous ferez d'abord une petite étude des ontologies existantes dans le domaine choisi. S'il y en a, en les analysant rapidement et en expliquant en quoi la vôtre sera différente.
2. **Construction d'une ontologie** : Construire votre ontologie en Protégé. Vous expliquerez, dans votre rapport, la méthode suivie pour cette construction. Les différents éléments (classes, propriétés) doivent être documentés dans Protégé. Il faut utiliser le maximum des différentes possibilités de représentation qu'offre OWL (CNS, connecteurs, restrictions, classes disjointes, etc). Ce sera un des critères d'évaluation du projet ainsi que la qualité de la documentation.
3. **Création des classes et des attributs** : Créer au moins 20 classes. Les classes seules ne fourniront pas assez d'information pour répondre aux questions de compétence. Après avoir défini quelques classes, il faudra décrire la structure interne des concepts, c'est-à-dire les propriétés et attributs. Vous devez créer au moins 10 attributs de différents types (nombres, string, énumérations, etc.).
4. **Création d'instances** : Construire plusieurs dizaines d'instances (minimum 30 instances) en complétant leurs propriétés qui lient ces instances entre elles. Elles permettront de tester la cohérence de votre ontologie sur un cas particulier mais un ensemble trop important vous ferait perdre trop de temps.

5. **Vérification de consistance** : En complément, lancer PELLET pour vérifier la consistance de votre ontologie.
6. **Requêtes SPARQL** : Construire quelques requêtes SPARQL (fédérées). Les tester dans Protégé ou à l'aide de Jena. Vous les stockerez à l'endroit adéquat et vous les donnerez dans le rapport.
7. **Point d'accès SPARQL** : Exposer un point d'accès SPARQL et créer un site Web simple contenant une interface graphique pour la saisie des requêtes et l'affichage des résultats.
8. **Exemples d'utilisation** : Donner quelques exemples d'utilisation potentielle en termes de Web sémantique.
9. **Livrable** L'ensemble des documents à fournir doit être contenu dans un même dossier portant votre nom. Ce dossier doit ensuite être compressé et envoyé via la plateforme pédagogique Moodle de notre faculté. Le dossier doit comporter :
  - **Le fichier OWL** : Un fichier OWL contenant l'ontologie construite à l'aide de Protégé. Cette ontologie doit être documentée et peuplée.
  - **Le rapport** : Un rapport qui explique l'ensemble du projet, les choix faits, le cœur de l'ontologie, les instances peuplant l'ontologie, le processus de création du point d'accès SPARQL, les requêtes SPARQL testées et les résultats, les exemples d'utilisation potentielle en termes de Web sémantique

### Etape 3 : Soutenance

1. **Fichier PowerPoint**: Le projet devra être soutenu oralement. Vous devrez créer un fichier PowerPoint ou équivalent contenant les diapositives de présentation et le sauvegarder en format PDF.
2. **Présentation** : La durée de la présentation est de 10 minutes + 5 minutes de questions & réponses. Durant la présentation, vous devrez reprendre succinctement chaque partie du rapport rendu et faire une démonstration en direct de votre point d'accès SPARQL (par exemple, aller sur le site du point d'accès, exécuter des requêtes SPARQL, visualiser des résultats).

### ÉVALUATION DU PROJET

- Mode d'envoi : **via la plateforme MOODLE** :  
<https://moodle-lettres.sorbonne-universite.fr>
- Première étape – Description du sujet :
  - Date : **lundi 20 mars 2023**
  - Livrable : rapport en PDF décrivant le domaine choisi
- Deuxième étape – Réalisation de l'ontologie :
  - Date : **lundi 17 avril 2023**
  - Livrable : rapport en PDF détaillant le travail effectué + fichier OWL
- Troisième étape – Soutenance :
  - Date : **lundi 17 avril 2023**
  - Livrable : présentation en PDF

Tout retard entraînera un zéro à la note de l'étape.