



UNIVERSITÉ DE SFAX
FACULTÉ DES SCIENCES DE SFAX
DÉPARTEMENT D'INFORMATIQUE

RAPPORT DE PROJET

Web Sémantique

Ontologie du domaine Éducation - Data Engineering

Encadré par : Mme Corine Zayani

Réalisé par :

- Mohamed Amine Salhi
- Oussema Zgha

Année universitaire 2024–2025

Remerciements

Nous exprimons nos sincères remerciements à **Madame Corine Zayani** pour son encadrement précieux durant le semestre 2, dans le cadre de l'enseignement de la matière *Web Sémantique*. Ses conseils et son accompagnement ont été essentiels à la réalisation de ce projet. Grâce à sa pédagogie et à sa disponibilité, nous avons pu comprendre et appliquer les concepts avancés du Web sémantique dans un cas concret et technique.

Table des matières

1	Objectif du projet	3
2	Introduction	3
3	Phase 1 : Choix du domaine et concepts clés	3
4	Phase 2 : Démarche de modélisation RDF/RDFS	4
4.1	Technologies employées	4
4.2	Application dans notre projet	4
4.3	Intérêt de la démarche RDF	4
5	Phase 2.5 : Schéma RDFS généré avec Protégé	5
6	Phase 3 : Requêtes SPARQL	6
6.1	Objectif des requêtes	6
6.2	Requête 1 : Apprenants, projets, technologies	6
6.3	Requête 2 : Modules, compétences, ressources	6
6.4	Requête 3 : Enseignants, cours, modules	6
6.5	Requête 4 : Apprenants, cours, modules	6
7	Phase 5 : Règles SWRL	7
8	Utilisation pratique de l'ontologie (Python)	7
8.1	Objectif du script Python	7
8.2	Fonctionnement du code Python	7
8.3	Résultat et impact	8
9	Conclusion	8

1 Objectif du projet

Ce projet vise à modéliser une ontologie dans le domaine de l'éducation en utilisant les technologies du Web sémantique : RDF, RDFS, OWL, SPARQL et SWRL.

2 Introduction

Le Web sémantique est une extension du Web actuel qui vise à rendre les données compréhensibles et exploitables par des machines. L'objectif est de permettre aux systèmes informatiques de raisonner automatiquement sur les données disponibles, en facilitant l'interconnexion et l'interrogation intelligente de l'information.

Dans ce contexte, ce projet s'inscrit dans une démarche de modélisation d'une ontologie liée à l'Éducation, plus précisément à la formation en Data Engineering. Notre objectif était de construire un modèle logique représentant les entités clés du domaine (apprenants, modules, compétences, ressources, etc.) et leurs relations, pour ensuite permettre l'interrogation intelligente via SPARQL et la déduction de nouvelles connaissances grâce à SWRL.

3 Phase 1 : Choix du domaine et concepts clés

Nous avons choisi le domaine de l'**Éducation orientée Data Engineering** pour plusieurs raisons :

- Il est directement lié à notre formation et donc facilement compréhensible.
- Il comporte de nombreux concepts clés interconnectés (cours, modules, compétences, projets, etc.).
- Il présente un grand intérêt pédagogique pour la recommandation de parcours ou le suivi d'apprenants.

Les concepts retenus pour notre ontologie incluent :

- **Learner** : un apprenant inscrit à une formation.
- **Instructor** : un enseignant responsable de cours.
- **Course, Module, Skill** : composantes académiques.
- **Project, Technology, Assessment, LearningResource** : activités et outils.

4 Phase 2 : Démarche de modélisation RDF/RDFS

La modélisation de notre ontologie a suivi les principes fondamentaux du Web sémantique :

4.1 Technologies employées

- **XML** : syntaxe d’encodage des données de base.
- **RDF (Resource Description Framework)** : pour représenter les données sous forme de triplets (sujet, prédicat, objet).
- **RDFS (RDF Schema)** : pour définir les classes et propriétés.
- **OWL (Web Ontology Language)** : pour ajouter des contraintes logiques (sous-classes, restrictions, `equivalentClass`, etc.).

4.2 Application dans notre projet

- Définition des classes comme `Learner`, `Instructor`, `Project`...
- Propriétés entre entités : `producesProject`, `usesTechnology`, `coversSkill`...
- Création des sous-classes (`Quiz`, `Exam`, `ProgrammingModule`) via RDFS.
- Utilisation des namespaces standards (`rdf`, `rdfs`, `xsd`, `foaf`, `dc`, `owl`).
- Enregistrement du fichier RDF dans Protégé, export au format RDF/XML.

4.3 Intérêt de la démarche RDF

Contrairement à une base relationnelle rigide, cette démarche permet :

- Une meilleure interopérabilité des données entre applications.
- Une sémantique explicite facilitant les inférences.
- Une extensibilité aisée pour ajouter de nouveaux concepts.

5 Phase 2.5 : Schéma RDFS généré avec Protégé

Pour visualiser graphiquement les relations entre les entités définies dans notre ontologie, nous avons utilisé l'outil Protégé, un éditeur open-source développé par l'université de Stanford. Cet outil nous a permis de :

- définir les classes, les propriétés et les restrictions de l'ontologie,
- vérifier la cohérence logique de notre modèle RDF/RDFS,
- exporter automatiquement un schéma visuel représentant les relations entre concepts.



Figure : Vue RDFS de l'ontologie depuis Protégé

Ce schéma aide à mieux comprendre la structure logique de notre ontologie avant de l'étendre avec OWL et de l'interroger avec SPARQL. Il confirme la hiérarchie des classes et la connectivité des propriétés entre les différentes entités pédagogiques.

6 Phase 3 : Requêtes SPARQL

6.1 Objectif des requêtes

Les requêtes SPARQL permettent d'interroger l'ontologie pour en extraire des relations sémantiques utiles à l'analyse pédagogique.

6.2 Requête 1 : Apprenants, projets, technologies

Objectif : lier les apprenants à leurs projets et les technologies utilisées.

```
SELECT ?learner ?project ?technology
WHERE {
  ?learner rdf:type :Learner .
  ?learner :producesProject ?project .
  ?project :usesTechnology ?technology .
}
```

6.3 Requête 2 : Modules, compétences, ressources

```
SELECT ?module ?skill ?resource
WHERE {
  ?module :coversSkill ?skill .
  ?module :hasResource ?resource .
  ?module rdf:type :ProgrammingModule .
}
```

6.4 Requête 3 : Enseignants, cours, modules

```
SELECT ?instructor ?course ?module
WHERE {
  ?instructor :teaches ?course .
  ?course :includesModule ?module .
  ?instructor rdf:type :Instructor .
}
```

6.5 Requête 4 : Apprenants, cours, modules

```
SELECT ?learner ?course ?module
WHERE {
  ?learner :enrolledIn ?course .
  ?course :includesModule ?module .
  ?learner rdf:type :Learner .
}
```

7 Phase 5 : Règles SWRL

Les règles SWRL enrichissent l'ontologie en ajoutant des inférences. Voici les règles implémentées :

- **S1** : Si un apprenant suit un module de programmation, alors il développe la compétence ProgrammingSkill.
- **S2** : Si un apprenant suit un module Big Data, alors il développe la compétence BigDataSkill.
- **S3** : Tout module qui développe BigDataSkill est un BigDataModule.
- **S4** : Si un cours contient un module couvrant une compétence, le cours couvre aussi cette compétence.

```
Rule S1:
Learner(?x) ^ hasTaken(?x, ?y) ^ ProgrammingModule(?y)
-> developsSkill(?x, ProgrammingSkill)

Rule S2:
Learner(?x) ^ hasTaken(?x, ?y) ^ BigDataModule(?y)
-> developsSkill(?x, BigDataSkill)

Rule S3:
Module(?x) ^ developsSkill(?x, BigDataSkill)
-> BigDataModule(?x)

Rule S4:
includesModule(?c, ?m) ^ coversSkill(?m, ?s)
-> coversSkill(?c, ?s)
```

8 Utilisation pratique de l'ontologie (Python)

L'utilisation d'une ontologie RDF/OWL ne se limite pas à sa création dans Protégé. Dans notre projet, nous avons cherché à exploiter cette base de connaissances à travers un script Python. Cela démontre la valeur ajoutée du Web sémantique lorsqu'il est connecté à des outils de traitement automatisé.

8.1 Objectif du script Python

Notre script Python a deux objectifs principaux :

- Interroger dynamiquement l'ontologie à l'aide de requêtes SPARQL,
- Générer automatiquement un **CV académique** pour chaque apprenant à partir des connaissances inférées (projets réalisés, compétences acquises, technologies utilisées).

Les résultats sont ensuite enregistrés dans un fichier Excel grâce à **pandas**, offrant une visualisation immédiate et exploitable des profils d'étudiants.

8.2 Fonctionnement du code Python


```

from rdflib import Graph
import pandas as pd

# Chargement de l'ontologie
onto = Graph()
onto.parse("ontologie.rdf", format="xml")

# SPARQL pour g nerer CV par apprenant
query = """
PREFIX : <http://www.education-ontology.org/education-data-engineering#>
SELECT ?learner ?project ?technology WHERE {
    ?learner a :Learner .
    ?learner :producesProject ?project .
    ?project :usesTechnology ?technology .
}
"""

# Traitement et export Excel
results = [(str(r.learner), str(r.project), str(r.technology)) for r in
    onto.query(query)]
df = pd.DataFrame(results, columns=["Apprenant", "Projet", "Technologie"
    ])
df.to_excel("cv_etudiants.xlsx", index=False)

```

8.3 R sultat et impact

Cette application illustre parfaitement l'utilit  d'une ontologie dans un syst me p dagogique. Elle permet de g n rer automatiquement des profils d taill s d'apprenants, exploitables par les enseignants, les plateformes  ducatives ou les responsables de stage. Elle constitue une passerelle concr te entre la mod lisation s mantique et les besoins m tiers.

9 Conclusion

Notre projet a mis en pratique l'ensemble des briques du Web s mantique :

- RDF/RDFS pour mod liser les donn es  ducatives,
- OWL pour structurer logiquement les concepts et relations,
- SPARQL pour les interroger intelligemment,
- SWRL pour enrichir la base de connaissances par inf rence,
- et Python pour connecter ces connaissances   des usages concrets.

Ce travail valide la faisabilit  et la puissance des ontologies dans un domaine complexe comme la formation en ing nierie des donn es, en fournissant une base extensible, interop rable, et exploitable automatiquement.