VIETNAM NATIONAL UNIVERSITY (VNU) INSTITUT FRANCOPHONE INTERNATIONAL (IFI)





Option : Systèmes Intelligents et Multimédia (SIM)

Promotion: XXI

Projet sur l'Ontologie et Web Sémantique

« Simulation d'un réseaux sociaux »

Auteurs:

MALLE Zoumana

MEDOU Daniel Magloire

RAKOTOARIVELO Anjara Nobby

Encadrant:

Monsieur TA Tuan Anh

année académique 2017-2018

TABLE DES MATIÈRES

1	Introduction							
	1.1	Contexte	3					
	1.2	Problématique	3					
	1.3	Objectifs et problème a résoudre	4					
	1.4	Plan de travail	4					
2	Analyse du sujet							
	2.1	Description du projet	5					
	2.2	Diagramme des cas d'utilisations	5					
	2.3	Description des cas d'utilisations	6					
	2.4	Termes techniques	7					
	2.5	Approches choisis	8					
3	Conception							
	3.1	Diagramme de classe système	8					
	3.2	Ontologie du système	9					
4	Implémentions							
	4.1	Environnement matériel	10					
	4.2	Environnement logiciel	10					
	4.3	Démarche adopter	11					
	4.4	Construction de la base de connaissance	12					
		4.4.1 Présentation Protégé OWL	12					
	4.5	Déploiement de la base de connaissance	13					
	4.6	Développement de l'application	14					
5	Pré	sentation de l'application	15					
6	Conclusion générale							
\mathbf{R}	éfére	ences	16					
	Arti	icles only	16					

LISTE DES FIGURES

1	Diagramme des cas d'utilisations	6							
2	Diagramme de classe système	9							
3	ontologie du système	10							
4	Processus de développement	12							
5	Présentation Protege OWL	12							
6		13							
7	Présentation de l'interface de JENA FUSEKI	13							
8	Démarrage serveur JENA FUSEKI	14							
9	Résultat d'exécution d'un requête SPARQL	14							
10	Présentation de l'application	15							
LISTE DES TABLEAUX									
$\frac{1}{2}$	Tableau des classes système	5 7							

LISTE DES ALGORITHMES

1 Introduction

Au terme de l'unité d'enseignement intitulée : ONTOLOGIE ET WEB SE-MANTIQUE, il a été soumis à notre étude un projet donc la réalisation dans notre cas devra être effectuée en trinôme. L'objectif général de ce dernier étant de simuler un réseau social tout en s'inspirant de l'une des technologies du web sémantique suivante RDF, OWL et XML. Dans le cadre de ce projet et de notre groupe en particulier, nous avons choisi d'implémenter une application de simulation d'un réseau social en utilisant l'approche du modèle RDF (Resource Description Framework). A la suite de ce rapport, il vous sera présenté de façon détaillée les différentes parties que constituent notre projet et programme et une partie expérimentations.

1.1 Contexte

L'influence grandissante des nouvelles technologies de l'information et de la communication à l'air des réseaux sociaux, ont littéralement changer notre vision sociale ou les barrières des liens sociaux peuvent être déterminer à travers d'algorithme structurer de données. L'explosion des fonctionnalités sociales au sein des applications du Web a favorisé le déploiement d'un panorama de médias sociaux permettant aux utilisateurs de librement contribuer, de se regrouper et d'interagir entre eux. La combinaison de divers moyens de publication et de socialisation permet de rapidement partager, recommander et propager l'information dans son réseau social, ainsi que de solliciter des réactions et de nouvelles contributions. Ces espaces partagés ont favorisé la création et le développement de communautés d'intérêts qui publient, filtrent et organisent de vastes répertoires de références dans leurs domaines, avec une impressionnante réactivité aux changements. Afin de reproduire les succès du Web dans la gestion d'information, de plus en plus de plates-formes sociales sont déployées dans des intranets d'entreprise. Cependant, l'avantage de ces plates-formes est fortement atténué lorsque le réseau social devient si grand que les informations pertinentes sont novées dans des flux continus de notifications. Organiser cette énorme quantité d'informations est l'un des défis majeurs du Web 2.0 afin de tirer pleinement partie des bénéfices de l'Entreprise 2.0, à savoir, l'utilisation des technologies du Web 2.0, tel que les blogs et les wikis, dans un intranet.

1.2 Problématique

L'analyse des réseaux sociaux propose des algorithmes de graphes complexes pour caractériser la structure d'un réseau social et ses positions stratégiques. Les technologies du Web Sémantique permettent de représenter et d'échanger les connaissances entre des applications distribuées sur le Web avec un modèle de graphes richement typés (RDF), un langage de requête (SPARQL) et des langages de description de modèles (RDFS et OWL). Le web change et, avec lui, changent les enjeux de l'éditorialisation. Ainsi, nous sommes passés du web statique des origines (1.0) au web participatif (2.0), puis au web sémantique (3.0) qui permet aux machines de comprendre la signification des données et de mieux les exploiter. Nous considérons que le passage au web sémantique est un enjeu majeur. Dans ce domaine, ce sont nos choix qui détermineront la structuration des connaissances dans le futur. Ce passage n'est pas neutre et

comporte une série de questionnements politiques, philosophiques, économiques, sociaux et techniques.

1.3 Objectifs et problème a résoudre

Les interactions des utilisateurs au travers des usages du web 2.0 amènent la communauté scientifique à réfléchir sur les moyens de capter ces usages pour y appliquer les techniques d'analyse des réseaux sociaux. Les applications bien connues à l'origine de l'émergence du web 2.0 sont les blogs, les wikis (ex : wikipedia), les services de social bookmarking (ex : del.ico.us), les sites de partages de médias (ex : youtube, flickr) et bien sûr les sites de réseaux sociaux (ex : facebook, LinkedIn). Ces applications ont considérablement accru la participation, les interactions et le partage entre les utilisateurs du web. L'analyse et la compréhension de tels réseaux sociaux suscitent de vifs intérêts au sein de plusieurs communautés scientifiques. Le web sémantique fournit des formalismes pour la représentation sémantique des personnes et de leurs usages sur le web. Ainsi l'objectif du-dit programme est fondé essentiellement sur l'orientation d'informations principalement afin d'effectuer une simulation à travers une base de connaissance de déterminer les potentiels liens sociales dérivant de l'approche des réseaux sociaux. le système permet d'aider un utilisateur entre autres :

- - Retracer l'arbre généalogique sociale d'une personne donnée
- - Détermination des liens de parentés directe ou indirecte de n'importe quelle personne contenue dans la base de connaissance sociale pré-établie.
- -Web social plus structuré et organisé, inter-opérable grâce aux ontologies et RDF
- - faciliter la description, l'échange et la portabilité des données
- -Donner plus de sens à l'activité des utilisateurs dans les réseaux sociaux

1.4 Plan de travail

2 Analyse du sujet

Il existe plusieurs moyens pour résoudre notre sujet, mais comme déjà citer ci-haut dans l'introduction, le meilleur moyens pour concevoir notre système est d'utiliser l'approche sémantique et ontologique. L'arbre ontologique et l'engagement sémantique qu'il explicite mettent à disposition des primitives. Il est alors possible de représenter formellement des connaissances puisque la modélisation ontologique a dégagé ce que la formalisation des connaissances présuppose comme pré requis. Il devient possible de définir une sémantique formelle pour les concepts et de retrouver les propriétés auxquelles on est habitué en représentation des connaissances. Il devient possible de définir un terme par une sémantique référentielle, c'est-à-dire une sémantique formelle. les concepts formels vérifient les relations d'identité unissant les concepts sémantiques. L'engagement ontologique permet ainsi de définir ce que nous appelons une ontologie formelle ou ontologie référentielle. Une ontologie référentielle est constituée de prédicats formels pourvus d'une sémantique référentielle [Charlet 2002].

2.1 Description du projet

Le projet est constitué de 7 entités qui sont :

Personne: une personne dans le réseaux sociaux

Homme : les personnes de sexe masculin
Femme : les personnes de sexe féminin
Ville : la ville où habite une personne
Pays : le pays où se situe la ville

École : l'école où les étudiants étudie

Entreprise: l'entreprise où les personnes travail

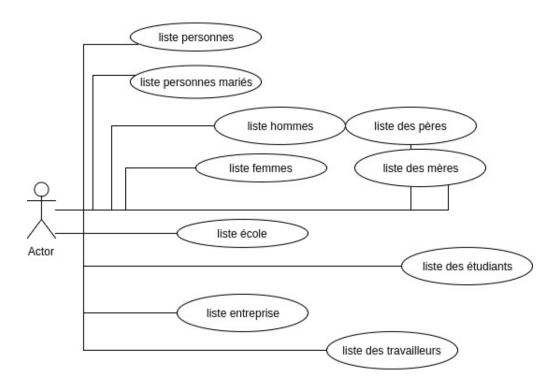
(Table 1).

Classe	Propriétés d'objet	Propriété des données	Cardinalités	Type de valeur
Personne	Travail	$nom_entreprise$	(1,1)	String
	étudie	nom_ecole	(1,1)	String
	habite	nom_ville	(1,1)	String
Homme	marie		(1,1)	
	père		(1,1)	
Femme	marie		(1,1)	
	mère		(1,1)	
Entreprise		$nom_entreprise$	(1,1)	String
Ecole		nom_ecole	(1,1)	String
Ville		nom_ville	(1,1)	String
Pays		nom_pays	(1,1)	String

Table 1 – Tableau des classes système

2.2 Diagramme des cas d'utilisations

Le Diagramme va nous permettre d'énumérer les futures fonctionnalités de l'application que nous allons développer, représenter par la (FIGURE 1)



 $FIGURE\ 1-Diagramme\ des\ cas\ d'utilisations$

2.3 Description des cas d'utilisations

(TABLEAU 2)

Numéro	Nomination	Description
Cas 1	Lister personnes	Lister tous les personnes enre-
		gistrer dans la base de connais-
		sance.
Cas 2	Lister hommes	Lister tous les hommes enregis-
		trer dans la base de connais-
		sance.
Cas 3	Lister femmes	Lister tous les femmes enregis-
		trer dans la base de connais-
		sance.
Cas 4	Lister hommes mariés	Lister tous les hommes mariés
		enregistrer dans la base de
		connaissance.
Cas 5	Lister femmes mariés	Lister tous les femmes mariés en-
		registrer dans la base de connais-
		sance.
Cas 6	Lister écoles	Lister tous les écoles enregistrer
		dans la base de connaissance.
Cas 7	Lister entreprises	Lister tous les entreprises enre-
		gistrer dans la base de connais-
		sance.
Cas 8	Lister pères	Lister tous les pères enregistrer
		dans la base de connaissance.
Cas 9	Lister mères	Lister tous les mères enregistrer
		dans la base de connaissance.
Cas 10	Lister étudiants	Lister tous les étudiants enre-
		gistrer dans la base de connais-
		sance.
Cas 11	Lister travailleurs	Lister tous les travailleurs enre-
		gistrer dans la base de connais-
		sance.

Table 2 – Description des cas utilisations

2.4 Termes techniques

Ontologie : A l'origine domaine philosophique de la « science de l'être en tant qu'être », une ontologie est, selon l'entendement du Web sémantique, un ensemble structuré de savoirs. Une ontologie définit les termes employés pour décrire et représenter un domaine de connaissance.

OWL: Web Ontology Langage

Métadonnée : Une métadonnée est une information permettant de décrire une autre information, quels qu'en soient la nature et le support.

Jena Framework : pour le web sémantique en Java, Jena fournit un environnement permettant de travailler avec RDF, RDFS et OWL.

HTML: HyperText Markup language

Protégé : Editeur d'ontologies et framework de gestion des connaissances, développé en open-source au sein de l'université de médecine de Stanford.

RDF : Ressource Description Framework, permet de présenter des données et des métadonnées.

RDFS: RDF Schema, permet de définir des vocabulaires RDF.

XLM: eXtensible Markup Language.

Littéral : une chaîne de caractères qui peut être la valeur d'une propriété.

Objet : la partie d'un triplet qui est la valeur de la déclaration.

Prédicat : la partie de la propriété dans un triplet.

Propriété: une propriété est un attribut d'une ressource.

Ressource: certaines entités. Elle pourrait être une ressource Web comme une page Web, ou une chose physique concrète comme un arbre ou une voiture. Elle pourrait être une idée abstraite comme les échecs ou le football. Les ressources sont nommées par une URI.

Déclaration : un arc dans un modèle RDF, normalement interprété comme un fait.

Sujet: la ressource qui est la source d'un arc dans un modèle RDF.

Triplet : une structure contenant un sujet, un prédicat et un objet. Un autre terme pour une déclaration.

2.5 Approches choisis

Dans le cadre de notre projet, l'approche donc nous avons jetté notre dévolu est le RDF. Ayant comme définition de son sigle en anglais qui est Resource Description Framework qui trouve sa définition en français comme étant un standard (techniquement une recommandation du W3C) pour la description de ressources. RDF est, un modèle nous permettant de décrire des métadonnées pour les ressources. Ce pendant dans la phase théorique c'est à dire cours du module, Nous nous concentrons sur trois langages d'ontologie qui ont été proposées pour décrire les ressources Web. Ces derniers sont RDF, RDFS et OWL. Nous considérons d'abord la langue RDF, une langue pour exprimer les faits (se concentrant principalement sur la base de données). Les deux autres langues permettent RDF dans les faits limitant les domaines d'application particuliers : RDFS est assez simple, alors que OWL est beaucoup plus riche. RDF (Resource Description Framework) fournit un langage simple pour décrire les annotations sur les ressources Web identifiées par URIs. Ce sont des faits. Les contraintes sur ces faits dans des domaines particuliers seront indiqués dans RDFS ou OWL. Au vue de cette importance simpliste qu'est RDF, et son importance par rapport aux deux autres (RDFS et OWL), nous avons choisi de par sont apport au autres langages de mettre en pratique notre projet avec le langage RDF.

3 Conception

3.1 Diagramme de classe système

Le diagramme de classe système de notre application est représenté par la $(FIGURE\ 2)$

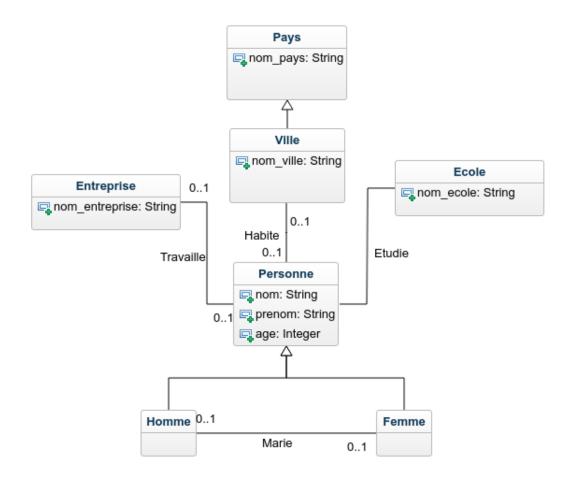


FIGURE 2 – Diagramme de classe système

3.2 Ontologie du système

(FIGURE 3)

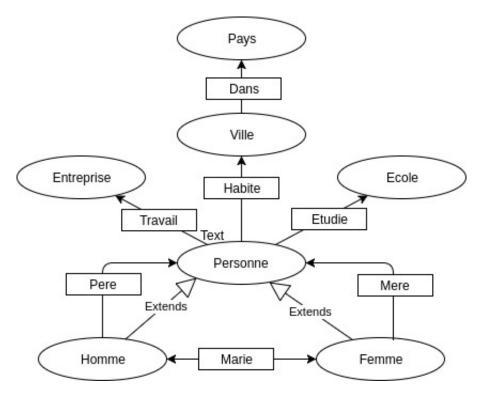


FIGURE 3 – ontologie du système

4 Implémentions

4.1 Environnement matériel

Pour la conception et les expérimentations de la simulation, nous avons utilisé un ordinateur portable « ASUS » avec les caractéristiques suivantes :

 $\mathbf{Processeur}: \, \mathrm{Intel}(R) \,\, \mathrm{CoreT} \,\, \mathrm{M} \,\, \mathrm{i5\text{-}M370} \,\, @2.4 \,\, \mathrm{GHZ}$

RAM: 4.00 Go **OS**: UBUNTU 16.4

4.2 Environnement logiciel

Pour la réalisation de notre projet, les outils et frameworks cité ci-dessous ont été d'une importance capital. Nous avons entre autre :

Editeur d'ontologies Protégé: C'est un logiciel gratuit (JAVA), plateforme open-source ui fournit une suite d'outils permettant la construction des bases de connaissances et des ontologies. Nous pouvons ajouter après exploration de ce dernier inclut des plusieurs plugins nous permettant de manipuler et de représenter des ontologies dans différents formats.

Framework JENA: Jena est un framework Java pour la construction d'applications Web sémantiques. Il fournit un environnement programmatique pour RDF, RDFS et OWL, SPARQL et inclut un moteur d'inférence basé

sur des règles. En effet, Jena est une API Java qui peut est utilisé pour créer et manipuler des graphes RDF. Dans le cadre de notre projet, il nous a permit de réaliser le processus de récupération des informations contenues dans le fichier OWL généré par Protégé.

Framework JENA Fuseki Serveur: Apache Jena Fuseki est un serveur SPARQL. Ce dernier peut fonctionner en tant que service du système d'exploitation par exemple une application Java Web (fichier War), et en tant que serveur autonome. Il assure la sécurité (en utilisant Apache Shiro) et dispose d' une interface utilisateur pour la surveillance et l'administration serveur. Il peut être utilisé pour fournir le moteur de protocole pour d'autres systèmes de requête et de stockage RDF.

- RDF (Ressource Description Framework): Quand à lui n'est pas à proprement parler un langage. Il s'agit plutôt d'un modèle de données pour décrire des ressources sur le web. Ici, nous entendons par ressource, tout entité que nous voulons décrire sur le web mais qui n'est pas nécessairement accessible sur le web.
- XML: Le XML ou eXtensible Markup Language est un langage informatique de balisage générique. En d'autre terme, XML est un langage qui permet de décrire des données à l'aide de balises et de règles que l'on peut personnaliser.
- OWL: langage d'ontologies. OWL offre un moyen d'écrire des ontologies web. Ainsi, OWL offre aux machines une plus grande capacité d'interprétation du contenu web que RDF et RDFS, grace à un vocabulaire plus large et à une vraie sémantique formelle. Ce dernier nous offre trois sous langages qui sont entre autres; OWL Lite est le sous langage de OWL le plus simple, OWL DL est plus complexe que OWL Lite, OWL Full est la version la plus complexe d'OWL. Ce pendant, il existe entre ces trois sous langage une dépendance de nature hiérarchique: toute ontologie OWL Lite valide est également une ontologie OWL DL valide, et toute ontologie OWL DL valide est également une ontologie OWL Full valide.

Langage de programmation: JAVA

Environnement de développement (IDE) : eclipse

4.3 Démarche adopter

Le processus de développement de notre application se subdivise en 3 parties représente par la (FIGURE 4).

(FIGURE 4)

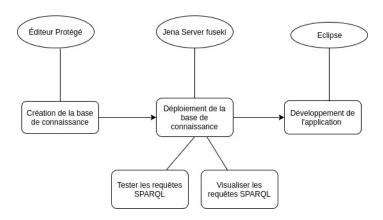


FIGURE 4 – Processus de développement

4.4 Construction de la base de connaissance

Comme déjà mentionner dans la partie (4.2 Environnement logiciel), nous avons utilisé l'éditeur Protégé pour la modélisation et conception de notre base de connaissance.

4.4.1 Présentation Protégé OWL

Présentation de l'interface d'accueil de Protégé (FIGURE 5).

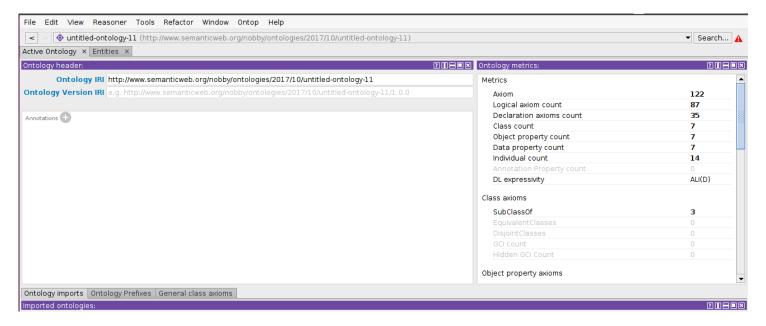


FIGURE 5 – Présentation Protege OWL

Déploiement de la base de connaissance

Après voir mis en place la base de connaissance, il faut déployer cette base. Ainsi nous avons utilise le serveur **JENA FUSEKI**.

Processus de déploiement du JENA FUSEKI :

- 1. Entrer dans le dossier contenant le serveur,
- 2. Démarrage du serveur avec la commande suivante :

```
./fuseki-server ---update ---mem /ds
```

Le serveur démarre sur le port 3030 ensuite, on obtient la (FIGURE 6) suivante.

```
nobby@nobby-pc: ~/workspace/websemantique/apache-jena-fuseki-3.4.0
obby@nobby-pc:~/workspace/websemantique/apache-jena-fuseki-3.4.0$ ./fuseki-serv
r --update --mem /ds
2017-11-20 12:31:10]
                                                 INFO Dataset: in-memory
INFO Apache Jena Fuseki 3.4.0
INFO FUSEKI_HOME=/home/nobby/workspace/websema
                                Server
2017-11-20 12:31:11]
                                Server
2017-11-20 12:31:11 Config
 ique/apache-jena-fuseki-3.4
2017-11-20 12:31:11] Config
                                                          FUSEKI_BASE=/home/nobby/workspace/websema
                                                 INFO
 ique/apache-jena-fuseki-3.
2017-11-20 12:31:11] Config
                                                 un
                                                 INFO
                                                          Shiro file: file:///home/nobby/workspace/
                                                 INFO Site File: File:///Home/Hobby/Workspace/
3.4.0/run/shiro.ini
INFO Template file: templates/config-mem
INFO Register: /ds
INFO Started 2017/11/20 12:31:14 ICT on port 3
 bsemantique/apache-jena-fus
017-11-20 12:31:12] Config
017-11-20 12:31:13] Config
 017-11-20 12:31:14]
```

Figure 6 – Démarrage serveur JENA FUSEKI

3. connexion sur JENA FUSEKI : la connexion sur le serveur s'effectue en allant sur le l'url http://localhost:3030/ (FIGURE 7).



Datasets on this server

```
1 Use the following pages to perform actions or tasks on this server:
                             Dataset Run queries and modify datasets hosted by this server.
                                       Administer the datasets on this server, including adding datasets, uploading data and performing backups
                   Manage datasets
                                       Summary of commands and links to online documentation.
                                Help
```

FIGURE 7 – Présentation de l'interface de JENA FUSEKI

4. déploiement de la base de connaissance : Pour déployer la base, il faut

tous simplement aller dans le menu dataset ensuite upload file (la base à ajouter) (FIGURE 8).

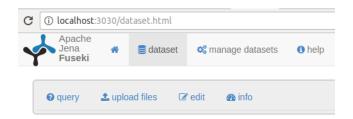


Figure 8 – Démarrage serveur JENA FUSEKI

5. tester et visualiser les requêtes SPARQL : la dernière étape consiste a construire et exécuté les requêtes SPARQL (FIGURE 9).

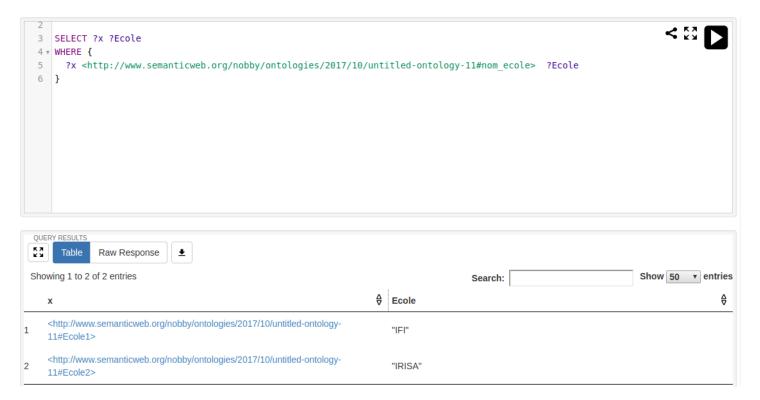


FIGURE 9 – Résultat d'exécution d'un requête SPARQL

4.6 Développement de l'application

Du coté application, nous avons un programme du coté client. C'est-à-dire on se connecte sur le serveur pour pouvoir traiter les informations voulus.

L'algorithme principale du programme est représenté par (ALGORITHME 4.6).

```
public static void execSelectAndProcess(String serviceURI, String
    query) {
    QueryExecution q = QueryExecutionFactory.sparqlService(
        serviceURI,
        query);
    ResultSet results = q.execSelect();

while (results.hasNext()) {
    QuerySolution soln = results.nextSolution();
    // assumes that you have an "?x" in your query
    RDFNode x = soln.get("x");
    System.out.println(x);
}
```

5 Présentation de l'application

La (FIGURE 10) nous donne un aperçu générale de notre application.

Figure 10 – Présentation de l'application

6 Conclusion générale

Arrivée au terme de notre projet donc l'objectif était lié au développement et la création d'une ontologie sur la simulation d'un réseau social, ceci dans le but de mettre en pratique la théorie de l'unité d'enseignement intitulée Ontologie et Web Sémantique. La mise en oeuvre de ce projet a amené l'équipe à effectuer un regroupement de nombreux concepts ainsi que les liens entre ces concepts (classes et sous classes), des relations entre ces concepts, leurs propriétés et des instances. Dans la réalisation du projet qui a été d'un grande importance voir même dune importance capitale, nous avons développé un outil/application pour la gestion de notre ontologie, qui structure sous forme de hiérarchique les différentes relations, entre les classes et les sous classe, et servira aux utilisateurs comme outil de recherche et de traitement de l'information. Le choix du langage d'implémentation joue un rôle important dans la construction d'une application donnée. Une liste d'outils de construction d'ontologie est présentée. L'état de l'art que nous avons présenté concerne les notions liées aux ontologies ainsi que les langages et techniques qui permettent de représenter les ontologies utilisés dans le cadre d'un système d'informations. En utilisant le langage JAVA et la bibliothèque Jena, nous avons développé un outil de simulation d'un réseau social.

Références

Charlet, Jean (2002). "L'ingénierie des connaissances : développements, résultats et perspectives pour la gestion des connaissances médicales". In :

Articles only

Charlet, Jean (2002). "L'ingénierie des connaissances : développements, résultats et perspectives pour la gestion des connaissances médicales". In :