République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université de Larbi Tébessi -Tébessa



Mémoire de Master

Présenté à l'Université de Tébessa
Faculté des Sciences Exactes et Sciences de la Nature et de la Vie

Département de : **Mathématiques et Informatique**
Spécialité : **Informatique**
Option : **SI**

Par :

Mr: ABDERRAHMANE Faouzi
Mlle: BENAOUA Aouatef

Thème

Interrogation de l'Ontologie DBpedia par les mots-clés

Jury

Président	**Professeur**	**Laouar Med. Ridha**	**Université de Tébessa**
Rapporteur	**Docteur**	**Bourougaa Salima**	**Université de Tébessa**
Encadreur	**Docteur**	**Djeddai Ala**	**Université de Tébessa**

Remerciements

Nous remercions avant tout, Dieu de nous a prodiguée la force morale et physique et nous a permis d'achever ce travail.

Ce travail n'aurait jamais pu voir le jour sans la contribution et l'effort de Mr. DJEDDAI Ala (Docteur) à qui nous tenons à exprimer notre profonde gratitude pour avoir bien voulu accepter de diriger ce travail.

Nos respects et notre gratitude vont également à MR. LAOUAR Med. Ridha (Professeur) , pour avoir accepté de présider le jury de cette thèse.

Nos profonds remerciements à Mme BOUROUGAA Salima (Docteur) pour avoir accepté d'être membre du jury de cette thèse.

Nous remercions tous les enseignants du département d'informatique que nous respectons beaucoup.

Enfin, nous remercions toutes nos familles et nos amis.

Dédicaces

À la mémoire de mes très chers parents.
À ma femme et mes enfants.
A tous mes frères et mes sœurs.
À tous mes amis et mes collègues .
À tous ceux qui sont proches de mon
cœur et dont je n'ai pas cité les noms.
Je dédie ce modeste travail.

" ***A. Faouzi*** "

Á mes très chers parents
Á mes frères et leurs familles
Á mes Amis.
Je dédie ce modeste travail.

B . Aouatef

Résumé – Abstract

Résumé :

Le Web sémantique et les ontologies jouent un rôle pivot notamment dans le contexte de la recherche d'information dans le but de pallier aux insuffisances des systèmes existants.

La sémantique d'une requête d'un utilisateur exigeant en termes de précision est difficile à extraire par une machine a l'aide des techniques classique. Le défi d'interroger la base de connaissances DBpedia par les mots clé nous obliges d'adapter une approche de recherche fiable pour résoudre les problèmes d'ambigüités entre les termes et les concepts.

Une solution proposée formalise le problème de la construction des requêtes formelles meilleures en traduisant les requêtes mots-clés en requêtes SPARQL en utilisant les connaissances disponibles dans l'ontologie DBpédia .

Abstract :

The Semantic Web and ontologies play a role particular in the context the search for information with the aim of mitigating the inadequacies of the existing systems.

User semantics of a request requiring in terms of precision is difficult to extract by a machine has the classic help(assistant) of techniques. The challenge to question the knowledge base DBpedia by keywords oblige us to adapt an approach of reliable search(research) to solve the problems of ambiguities between the terms and the concepts.

A proposed solution formalizes the problem of the construction of the better formal requests by translating the requests keywords into requests SPARQL by using the available knowledge in the ontology DBpedia .

Sommaire

	Page
Introduction Générale	**1**
Chapitre 1: Principes Généraux	
1.1. Le Web Sémantique	3
1.1.1. Définition	3
1.1.2. Le World Wide Web Consortium (W3C)	5
Protocoles utilisés	5
a- le protocole HTTP	5
b- les U R I	5
c- le langage XML	5
1.1.3. Annotation sémantique et concept de métadonnée	6
a- Annotation sémantique	6
b- Concept de métadonnée	7
1.1.4. Les Ressources Sémantiques (“ Les Ontologies”)	8
a- Taxonomie	8
b- Thesaurus	8
c- L'ontologie	9
Type d’ontologie	10
L’utilité des Ontologies	10
Ontologies Majeures	10
Web de données liées (Linked Data)	12
Web de données ouvertes et liées (Linking Open Data)	12
1.1.5. L'ontologie Dbpedia	12
1.2. Les Langages des Ontologies	14
1.2.1. RDF	14
a- Les syntaxes de RDF	15
b- Les graphes conceptuels (GCs)	16
1.2.2. RDFS (RDF Schema)	17
a- Les Classes	17
b- Les propriétés	18
c- Les Individus	18
1.2.3. OWL	19
1.2.4. SPARQL	19
a- Une requête SELECT	20
b- Une requête ASK	20
c- Une requête CONSTRUCT	20
d- Une requête DESCRIBE	21

1.3. La recherche d'information	21
1.3.1. Introduction	21
1.3.2. Définition	22
1.3.3. Modèles de recherche d'informations	23
a- Le modèle booléen	23
b- Le modèle vectoriel	24
c- Le modèle probabiliste	25
1.3.4. La recherche sémantique d'information	25
a- Indexation sémantique	26
b- Traitement sémantique de la requête	27
c- Appariement sémantique	27
1.4. Conclusion	27
Chapitre 2: Etat De L'Art	
Introduction	28
2.1. Description	28
- Interrogation en mode simple	28
- Interrogation en mode complexe	29
2.2. Détail des l'approches existantes	29
2.2.1. Interrogation en mode simple : (Approches Coopératives)	29
a - Suggestion de requêtes	29
b - Interrogation par l'exemple	30
c - Explications sur les réponses d'une requête	30
d - Raffinement de la requête par l'utilisateur (Users refinement)	30
e - Redéfinition des requêtes	31
2.2.2 Interrogation en mode complexe: (Approches basées sur la structure des ressources)	31
2.3 . Les approches selon le degré de formalité	32
2.4 . Modèles de Moteurs sémantique existants	34
2.5 . Conclusion	35
Chapitre 3 : La Conception	
Introduction	36
3.1 - Description	36
3.2 - Présentation du problème	37
. Exemple (1)	38
. Prefixes (URIs)	39
3.3 - Etapes du système	40
3.3.1. Mappage des mots clés aux éléments de la B.C	40
Définition	40

Le détail	41
. Le mappage morphologique	42
. Le mappage sémantique	42
. Suite de l'exemple (1)	42
3.3.2 Construction des Graphes Requête	44
a - Calcul des graphes de Connexions	44
. Explorer les liens entre les éléments de la B.C	44
. Calcul des graphes de connexions	47
b- Mappage des graphes de Connexions aux graphes requête	49
3.3.3 Classement des requêtes	51
• Algorithme de Classement	52
3.4 - Conclusion	53
Chapitre 4 : Implémentation et Evaluation	
Introduction	54
4.1. L'Implémentation	54
4.1.1. Les outils et les langages utilisés	54
4.1.2. L'Implémentation	55
a- Edition de l'ontologie DBpedia	55
b- Exploitation du code OWL dans le programme Java	55
c- Exécution de l'application	56
. Interface principale	56
. Interface après recherche sur le mot clé « book »	57
. Interface après la recherche sur la requête « writer misérable book »	57
4.2. Evaluation	58
4.2.1. Métriques utilisées	58
a- La précision	58
b- Le rappel	58
4.3 . Conclusion	59
Conclusion Générale	60

Liste des Figures

N°	Titre du Figure	Page
Figure 1.1	La pile des standards du web sémantique	4
Figure 1.2	Les groupes de la pile des standards du WS.	4
Figure 1.3	Exemple d'annotation Sémantique basée ontologie	6
Figure 1.4	Exemple de Métadonnées d'un fichier PDF	8
Figure 1.5	Vocabulaires et richesse sémantique	9
Figure 1.6	Le web de données en Avril 2017	11
Figure 1.7	L'Ontologie Dbpedia	12
Figure 1.8	Wikipedia avec infobox pour créer DBpedia	13
Figure 1.9	Représentation d'un triplet RDF (sujet , relation , objet).	14
Figure 1.10	Ensemble de triplets RDF	15
Figure 1.11	Représentation sous forme de graphe de l'ensemble de triplets RDF	15
Figure 1.12	Document représenté en figure 6 au format RDF/XML.	16
Figure 1.13	Exemple de graphe conceptuel	16
Figure 1.14	Exemple de l' hiérarchie de classes	17
Figure 1.15	Exemple de l'hiérarchie de propriétés	18
Figure 1.16	Exemple d'instances des types Homme, Femme et Lieu .	19
Figure 1.17	Exemple de requête SPARQL SELECT	20
Figure 1.18	Résultat de la requête SPARQL SELECT	20
Figure 1.19	Exemple de requête SPARQL ASK.	20
Figure 1.20	Exemple de requête SPARQL CONSTRUCT	21
Figure 1.21	Résultat de la requête SPARQL CONSTRUCT	21
Figure 1.22	Exemple de requête SPARQL DESCRIBE	21
Figure 1.23	Résultat de la requête SPARQL DESCRIBE	21
Figure 1.24	Productivité de la recherche	22
Figure 1.25	Système de recherche d'information	23
Figure 1.26	Appartenance des termes d'une requête dans un document	24
Figure 1.27	Le modèle vectoriel	24
Figure 1.28	Système de recherche sémantique de l'information	26
Figure 2.1	L'interface du moteur Corese	34
Figure 2.2	L'interface du moteur Evi	35
Figure 2.3	L'interface du moteur wolframalpha	35
Figure 3.1	Architecture générale de l'approche	37
Figure 3.2	Le Graphe Conceptuel extrait de l'ontologie DBpedia	39
Figure 3.3	Calcule de similarité	41

Figure 3.4	Exemple de mapping	43
Figure 3.5	Algorithme (1) Exploration de la B.C	45
Figure 3.6	Algorithme (2) Passé	46
Figure 3.7	Graphe résultat d'application de l'algorithme d'exploration	47
Figure 3.8	Algorithme (3) calcule des graphes de connexions	48
Figure 3.9	Graphe des Connexions 1	48
Figure 3.10	Graphe des Connexions 2	49
Figure 3.11	Graphe de requête 1	50
Figure 3.12	Graphe de requête 2	50
Figure 3.13	La requête Sparql du Graphe Requête (1)	50
Figure 3.14	La requête Sparql du Graphe Requête (2)	50
Figure 3.15	Classement des requêtes	51
Figure 3.16	Algorithme (4) - Classement des résultats.	52
Figure 4 .1	Vue de Librairies utilisés sur JAVA Eclipse	56
Figure 4 .2	Interface Principale de l'application.	56
Figure 4 .3	Résultat de recherche sur la requête « book »	57
Figure 4 .4	Résultat de recherche sur la requête « writer misérable book »	57

Liste des Tableaux

N°	Titre du Tableau	Page
Tableau 1.1	L'évolution de version Anglaise de DBpedia	14
Tableau 2.1	Classification des approches par KAUFMANN E. & BERNSTEIN A	33
Tableau 3.1	Les valeurs des similarités trouvées entre les mots clés et les ressources	43
Tableau 3.2	Classement des requêtes formelles selon leurs scores	52
Tableau 4.1	Ensembles des requêtes mots-clés	58
Tableau 4.2	Résultats de calcule des Rappels et Précisions	59