

**Ministere de l'enseignement supérieur et de la recherche scientifique  
INSTITUT NATIONAL DE FORMATION EN INFORMATIQUE (I.N.I.)  
Oued-Smar Alger**

**Mémoire de fin d'études  
Pour l'obtention du diplôme d'ingénieur d'état en informatique  
Option : systèmes d'information**

**Thème :**

**Outil d'enrichissement d'ontologie pour l'amélioration du  
processus d'indexation dans la recherche d'information**

**Réalisé par :**

**ANNOUCHE KARIM  
BENHOCINE RACHID**

**Proposé par :**

**Mme D. Keddari  
Mme L. Boumellil**

**Organisme d'accueil C.E.R.I.S.T  
Centre de recherche sur l'information scientifique et technique**

**PROMOTION / 2005-2006**

## *Table des matières*

### Introduction générale

---

#### Partie I : Revue de littérature

---

### Chapitre 1 : web sémantique

1. Introduction .....	1
2. Un Bref historique du web .....	2
3. Lacunes du web actuel.....	3
4. Qu'est ce qu'un web sémantique .....	4
5. Du web actuel au web sémantique .....	5
6. Architecture du web sémantique .....	7
7. Concepts de base du web sémantique .....	8
7.1. Ontologie .....	8
7.2. Méta-données .....	8
7.3. Annotation .....	9
8. Principes du web sémantique .....	10
9. Acteurs du Web Sémantique .....	11
10. Quelques applications du web sémantique.....	12
10.1 Web sémantique d'une entreprise .....	12
10.2 Web sémantique Médical (WSM).....	12
10.2.1 Usages du Web sémantique en médecine.....	13
10.2.2 Quelques travaux plus représentatifs dans le WSM.....	13
11. Conclusion.....	14

### Chapitre 2 : Ontologies et recherche d'information

1. Introduction .....	15
2. Définitions d'ontologie.....	15
3. Composantes d'une ontologie .....	16
4. Typologies des ontologies .....	17
5. Rôles et objectifs des ontologies dans le web sémantique .....	18
6. Caractéristiques des ontologies .....	19
7. Langages de spécification d'ontologie .....	19
7.1 RDF et RDF Shéma.....	20
7.2 OIL .....	20
7.3 DAML+OIL .....	21
7.4 OWL.....	21
8. Les éditeurs d'ontologies.....	21
8.1 Protégé2000.....	22
8.2 OILD.....	22
8.3 OntoEdit .....	22
9. Cycle de vie d'une ontologie .....	23
10. Méthodes d'évaluation des ontologies .....	24
11. Domaines d'application des ontologies.....	25
12. Thésaurus.....	26
12.1 Définition .....	26
12.2 Rôles du thésaurus.....	27
12.3 Relations Sémantiques dans un thésaurus .....	27

12.4 Exemples de thésauri .....	28
13. Recherche d'information .....	30
13.1 Système de recherche d'information .....	30
13.2 Concepts clés de la recherche d'information (RI) .....	31
13.4 Problématique liée aux requêtes dans les SRI classiques .....	32
13.5 Utilisation des l'ontologies par un SRI .....	32
13.6 Recherche d'information basée sur les ontologies .....	33
13.7 Apports de l'ontologie dans le domaine de la RI .....	35
13.8 Evaluation d'un système de recherche d'information .....	36
14. Conclusion .....	37

### Chapitre 3: Enrichissement des ontologies

1. Introduction .....	38
2. Evolution des ontologies .....	38
3. Approches d'enrichissement des ontologies .....	39
3.1 Apprentissage d'ontologie à partir de textes .....	40
3.2 Apprentissage d'ontologie à partir de dictionnaire .....	41
3.3 Apprentissage d'ontologie à partir de bases de connaissances .....	41
3.4 Apprentissage d'ontologie à partir de schémas semi-structurés .....	42
3.5 Apprentissage d'ontologie à partir de schémas relationnels .....	42
4. Exemples de quelques méthodes d'enrichissement .....	43
5. Quelques techniques d'enrichissement d'ontologies .....	44
5.1 Technique d'enrichissement à partir du corpus .....	44
5.2 Enrichissement par interrogation du thésaurus WordNet .....	46
5.3 Enrichissement avec élagage de l'ontologie .....	49
5.4 Enrichissement coopératif des ontologies .....	50
5.4.1 Qu'est ce que l'enrichissement coopératif des ontologies ? .....	50
5.4.2 Gestion des utilisateurs .....	51
5.4.3 Différents approches pour l'enrichissement coopératif .....	51
5.5 Enrichissement par fusion des ontologies .....	53
5.5.1 Fusion des ontologies .....	53
5.5.2 Technique d'enrichissement par fusion .....	54
5.5.3 Exemples de méthodes de fusion et d'intégration d'ontologies .....	56
6. Conclusion .....	57

---

### Partie II: Conception du système

---

1. Introduction .....	58
2. Objectifs .....	59
3. La technique retenue pour l'enrichissement .....	60
4. Schéma fonctionnel de l'outil d'enrichissement .....	61
5. Description des composantes de l'outil .....	63
5.1 Un Thésaurus .....	63
5.2 Une Ontologie .....	64
5.3 Module d'extraction de concepts du thésaurus .....	64
5.4 Module de recherche de concepts dans l'ontologie .....	66
5.5 Module de Mise à Jour d'ontologie (MAJ) .....	67
5.6 Module de validation .....	69
5.7 Module de gestion des versions d'ontologie .....	70
6. Scénario du fonctionnement du système .....	71
7. Conclusion .....	74

---

**Partie III: Réalisation du système**


---

1. Introduction .....	75
2. Environnement de développement .....	75
2.1 Choix du langage de programmation .....	75
2.2 Choix du format de l'ontologie .....	76
2.3 Choix du domaine de l'application.....	76
3. Architecture de EnriOnto .....	78
4. Implémentation du système.....	79
4.1 L'ontologie .....	79
4.2 Le thésaurus.....	79
4.3 Module d'enrichissement .....	80
4.3.1 Module d'extraction de concepts.....	81
4.3.2 Module de Recherche de concepts .....	81
4.3.3 Module de MAJ de l'ontologie.....	82
4.4 Module de validation.....	83
5. Mise en oeuvre de EnriOnto.....	83
6. Tests sur le temps d'exécution pour l'enrichissement d'une ontologie .....	95
7. Evaluation de EnriOnto .....	Erreur ! Signet non défini.
7.1 Résultats de l'enrichissement .....	Erreur ! Signet non défini.
7.2 Tests avec MIRO .....	98
7.2.1 Tests sur l'indexation .....	98
7.2.2 Tests sur la Recherche .....	101
8. Conclusion.....	Erreur ! Signet non défini.

**Conclusion générale****Annexes****Bibliographie**

## Résumé

Les ontologies sont des nouvelles formes de contrôle intelligent de l'information. Elles présentent un savoir préalable requis pour un traitement systématique de l'information à des fins de navigation, de rappel, de précision, etc. Toutefois, les ontologies sont confrontées de façon continue à un problème d'évolution. En effet, ce projet s'inscrit dans le cadre de recherche d'information basée sur les ontologies, notamment l'enrichissement des ontologies pour l'amélioration du processus d'indexation dans la recherche d'information. Cet enrichissement d'ontologie, supervisé par un expert du domaine, utilise un processus basé sur un thésaurus afin de découvrir de nouveaux concepts et relations entre concepts. Par conséquent, l'indexation des pages, actuellement, rejetées par le processus d'indexation est faite.

**Mots- clés :** Web sémantique, Ontologie, Enrichissement d'ontologie, Thésaurus, Recherche d'information.

## Abstract

Ontologies are new forms of intelligent control of information. They present a previous knowledge required for a systematic processing of information for navigation, recall, precision, etc. However, ontologies are confronted continuously with a problem of evolution. Indeed, this project lies within the scope of search for information based on ontologies, in particular the ontology enrichment for the improvement of the indexing process in the search for information. This ontology enrichment, supervised by an expert in the field, uses a process based on a thesaurus in order to discover new concepts and concept interlinks. Consequently, the indexing of the pages, currently, rejected by the process of indexing is made.

**Key Words:** Semantic web, Ontology, Ontology enrichment, Thesaurus, information retrieval.