



**République Algérienne démocratique et populaire**  
**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**  
**Université des Sciences et de la Technologie Houari Boumediene U.S.T.H.B**

Faculté d'électronique et d'Informatique  
*Département d'Informatique*

**Projet de fin d'études pour l'obtention du diplôme D'ingénieur d'état**

**Option : Recherche d'Information**

---

***THEME :***

**La recherche d'information basée sur Le profil utilisateur et l'usage  
- Cas d'étude le domaine de la recherche scientifique du CERIST-**

---

**Sujet proposé par :**

**M<sup>lle</sup> Nour El Houda Boulkrinat**

**Présenté par :**

**Benkherfallah Nabila Sabrina**

**Haddadj Souhila**

**Soutenu le : 12/10/2011**

**Devant le jury composé de :**

**M<sup>me</sup> N.ABDAT                      Présidente**

**M<sup>r</sup> A.Saadi                              Membre**

**M<sup>me</sup> R.Amrouni                      Membre.**

**Binôme N° : 08/2011**

## **Résumé**

*L'accès personnalisé à l'information permet de délivrer à l'utilisateur, à partir d'un large volume d'informations hétérogènes, seuls les documents susceptibles d'être appropriés à ses critères de recherche (centres d'intérêt, préférences,...) dont on appelle profil.*

*La modélisation du profil utilisateur est la clé de réussite d'un processus d'accès à l'information. Pour cela, nous avons essayé, dans ce travail, d'intégrer le profil utilisateur et l'objectif de sa recherche (usage) dans le processus de recherche d'information afin d'améliorer la pertinence des résultats. Nous avons adopté pour une représentation multidimensionnelle du profil vu qu'il est constitué de différents types d'informations. En plus de l'acquisition explicite des informations du profil, nous avons opté pour l'acquisition implicite basée sur le comportement de l'utilisateur avec le système comme source de prédiction des usages de l'utilisateur et ses centres d'intérêts.*

**Mot clés :** système de recherche d'information, profil utilisateur, modélisation de profil, ontologie, centre d'intérêt, usage.

## **Abstract**

*The personalized access to information can deliver to the user, from a large volume of heterogeneous information, only documents that may be relevant to their search criteria (interests, preferences ...), which is called profile.*

*Modeling the user profile is the key to a successful process access to information. For this, we tried in this work, to integrate the user profile and the objective of his research (use) in the information retrieval process to improve the relevance of the results. We adopted a multidimensional representation of the profile as it consists of different types of information. In addition to the acquisition of explicit profile information, we opted for the acquisition based on implicit user behavior with the system as a source of prediction uses the user and their interests.*

**Key words:** System of search for information, user profile, modeling profile, ontology, term interests, use.

## *Remerciement*

*Louange à Allah l'unique et le tout puissant qui nous a donné la force pour accomplir ce travail.*

*Nos sincères remerciements à notre promotrice Mlle Nour El Houda Boulkrinat de nous avoir fait confiance, de nous avoir suivies tout au long de ce projet, pour sa patience et sa disponibilité.*

*Un grand remerciement aux membres du jury, Mme N. Abdal, Mr A. Saadi et Mme R. Amrouni pour nous avoir honorées en consentant à juger notre modeste travail.*

*Nous tenons à remercier Mme Samia Boulkrinat pour sa précieuse aide et pour le temps qu'elle nous a accordé.*

*Merci à Mr Samir Kachid pour son aide et ses conseils.*

*Merci également à Mr Ahmed Guessoum pour son attention et son orientation.*

*Merci aux ingénieurs de l'INI, Rabeh Guerbi et Mahmoud Fatoumi pour leur aide.*

*Nous remercions enfin tous ceux qui ont contribué à la réalisation de ce modeste travail.*

# DEDICACE

*Nous remercions Allah le tout puissant qui nous a donné le courage et la volonté pour réaliser notre projet de fin d'études.*

*Je dédie ce modeste travail*

*A ma Mère, A ma Mère, et A ma Mère,  
A mon Père,*

Aux être les plus chers que j'ai dans la vie ma mère et mon père pour l'amour et l'affection qu'ils m'ont données, durant toute ma vie, Qu'Allah vous bénisse, vous assiste, vous soit en aide ;

*A mon frère **Karim** et à **Nasser -Adine**,*

*A mes chères et adorables soeurs **Amina** et **Fairouz***

J'adresse tout d'abord mes remerciements à des personnes importantes dans ma vie, Aucun mot sur cette page ne saurait exprimer ce que je veux dire, que ce soit de près ou de loin. Un grand merci pour :

- Docteur **Lila** pour sa gentillesse, sa patience,
- Docteur **Soraya**, pour l'encouragement
- La sincère **Rabia** pour son aide,
- J'offre un grand merci à mon frère **Radwan** pour son encouragement
- A **Karim abd eldjber**
- A mes professeurs de l'**USTHB**, En particuliers «**Mr Bouzaghoub, Mme boukala, Mme Derbal, Mr Selmoune, Mr Ait zai**»,

Je remercie également mon binôme **Sabrina** avec qui j'ai partagé, une année de travail agréable.

Un remerciement spécial à notre promotrice **Mlle Nour El Houda Boulkrinat** pour ses conseils, sa confiance et son encouragement ; vous étiez vraiment plus qu'une promotrice pour nous....merci Houda.

Je remercie mes amies de l'**USTHB**, particulièrement **Khadidja** et **Asma, Dalila, karima, Djalila, Daouia, Fatima, Zineb, Amina, Nesrine, Souhila, Yasmine**...et pour toute étudiante de la 5<sup>ème</sup> année Informatique promo 2011.

A tous mes enseignants de l'**USTHB**.

A vous tous, je dis **MERCI**.

*Souhila*

## *Dédicace*

*A la mémoire de mon père et de mon frère Hamidou que je n'oublierai jamais.*

*A ma très chère maman qui m'est une source d'amour, de tendresse, de bonheur... et de tout ce qui est beau dans ma vie, aucun mot ne pourra exprimer ce que je te dois maman, je ne pourrai jamais te rendre tes obligations, j'espère que tu trouveras dans ce travail toute ma reconnaissance et mon amour. Qu'Allah te protège et te garde pour moi.*

*A mon frère Sidali qui m'a toujours soutenu et encouragé tout au long du parcours de mes études.*

*A ma petite nièce Rania, Inch'Allah qu'elle aura une très longue et belle vie au sein de ses parents.*

*A toi Faiza, tu m'as toujours été l'exemple d'ambition et de succès, merci ma sœur pour ton aide et pour ton soutien.*

*A ma très chère amie Kahina, avec qui j'ai mes plus beaux souvenirs inoubliables.*

*A ma meilleure amie et chère sœur Zineb qui a toujours été disponible pour m'aider et m'encourager. Ton absence m'a démontré combien tu compte pour moi.*

*A toi Hadjer, mon adorable amie d'enfance, merci pour ta présence, ton soutien, et pour tes conseils.*

*Hadjer, Zineb, vous allez me manquer énormément mes amours. Je vous souhaite beaucoup de succès dans le parcours que vous avez choisi. Qu'Allah vous garde pour vos chers et pour notre amitié, Je vous adore.*

*A Souhila, mon binôme qui a partagé avec moi ce travail.*

*A mes amies de Babez, avec qui j'ai vécu des moments inoubliables, Fatima, Daouia, Djalila, Amina et Meriem. Vous êtes adorables les filles. Que notre amitié dure inch'Allah jusqu'à la fin de notre vie.*

*A tous ceux qui me sont chers et je n'ai pas cité leurs noms.*

*Fidèlement*

*Sabrina qui vous aime tous.*



# Table des matières

INTRODUCTION GENERALE.....	1
----------------------------	---

## Chapitre I : LA RECHERCHE D'INFORMATION

1. INTRODUCTION.....	3
2. DEFINITION DE LA RECHERCHE D'INFORMATION.....	3
3. NOTION DE BASE DE LA RI.....	3
3.1 LA REQUETE DE L'UTILISATEUR.....	4
3.2 LE DOCUMENT.....	4
3.3 LES SYSTEMES DE RECHERCHE D'INFORMATION (SRI).....	4
3.3.1 Définition d'un système de recherche d'information.....	4
3.3.2 Les outils de la recherche d'information.....	4
3.4 NOTION DE PERTINENCE.....	6
4. LE PROCESSUS DE RECHERCHE D'INFORMATION.....	7
4.1 L'INDEXATION.....	8
4.1.1 Modes d'indexation.....	9
4.1.1.1 Indexation manuelle.....	9
4.1.1.2 Indexation automatique.....	9
4.1.1.3 Indexation semi-automatique.....	9
4.1.2 Les étapes du processus d'indexation.....	9
4.1.2.1 Analyse lexicale.....	9
4.1.2.2 Élimination des mots vides.....	9
4.1.2.3 Lemmatisation.....	9
4.1.2.4 Pondération des termes.....	10
4.2 L'INTERROGATION.....	11
4.3 LA FONCTION D'APPARIEMENT DOCUMENT-REQUETE.....	11
4.4 LA REFORMULATION DE LA REQUETE.....	12
5. LES MODELES DE LA RI.....	12
5.1 LES MODELES ENSEMBLISTES.....	12
5.1.1 Le modèle booléen.....	12
5.2 LES MODELES ALGEBRIQUES.....	13
5.2.1 Le modèle vectoriel.....	13
5.3 LES MODELES PROBABILISTES.....	14
5.3.1 Le modèle probabiliste.....	14
5.4 AVANTAGES ET INCONVENIENTS DES MODELES.....	15
6. EVALUATION DES SYSTEMES DE RECHERCHE D'INFORMATION.....	16
6.1 MESURES D'EVALUATION.....	16
6.1.1 Le rappel.....	17
6.1.2 La précision.....	17
6.2 AUTRES MESURES D'EVALUATION.....	18
6.2.1 La Moyenne harmonique.....	18
6.2.2 La mesure E.....	18
7. LES ONTOLOGIES ET LA RECHERCHE D'INFORMATION.....	19
7.1 DEFINITIONS.....	19
7.2 CARACTERISTIQUES DES ONTOLOGIES.....	19
8. LES ONTOLOGIES ET LES SYSTEMES DE RECHERCHE D'INFORMATION.....	20
8.1 PRINCIPE D'UTILISATION DES ONTOLOGIES PAR UN SRI.....	20
9. DE LA RI CLASSIQUE A LA RI ADAPTATIVE.....	21
10. LA RI ADAPTATIVE.....	22
10.1 REFORMULATION DE REQUETES.....	22
10.2 DESAMBIGUÏSATION DU SENS DES MOTS DE LA REQUETE.....	23

10.3	LIMITATIONS DE LA RI ADAPTATIVE .....	23
11.	CONCLUSION .....	24

## **Chapitre II : PROFIL UTILISATEUR ET USAGE**

1.	INTRODUCTION.....	25
2.	NOTION DE PROFIL.....	25
3.	DOMAINE D'UTILISATION INFORMATIQUE DES PROFILS .....	26
4.	RECHERCHE D'INFORMATION ET LE PROFIL UTILISATEUR.....	26
5.	CARACTERISTIQUES DU PROFIL UTILISATEUR.....	27
5.1	TRAITS EXPLICITES.....	27
a.	<i>Connaissances</i> .....	27
b.	<i>Objectifs</i> .....	27
c.	<i>Références ou background et expérience</i> .....	27
d.	<i>Préférences</i> .....	27
e.	<i>Choix ou les centres d'intérêts</i> .....	28
f.	<i>Traits individuels</i> .....	28
5.2	TRAITS IMPLICITES DU PROFIL UTILISATEUR.....	28
6.	MODELISATION DU PROFIL UTILISATEUR.....	28
6.1	DONNEES PERSONNELLES .....	28
6.2	CENTRES D'INTERET .....	28
6.3	ONTOLOGIE DU DOMAINE .....	29
6.4	QUALITE ATTENDUE DES RESULTATS DELIVRES.....	29
6.5	CUSTOMISATION.....	29
6.6	SECURITE ET CONFIDENTIALITE .....	29
6.7	RETOUR DE PREFERENCES.....	29
6.8	INFORMATIONS DIVERSES .....	29
7.	ACQUISITION DU PROFIL UTILISATEUR.....	30
7.1	ACQUISITION EXPLICITE .....	30
7.2	ACQUISITION IMPLICITE .....	30
8.	INDICATEURS POUR LA DETECTION DES CENTRES D'INTERET .....	31
9.	INTEGRATION DU PROFIL UTILISATEUR DANS UN SRI.....	31
9.1	GESTION DES PROFILS.....	32
9.1.1	<i>Représentation du profil</i> .....	32
9.1.1.1	Représentation vectorielle .....	33
9.1.1.2	Représentation sémantique .....	33
9.1.1.3	Représentation connexionniste .....	33
9.1.1.4	Représentation multidimensionnelle.....	33
9.1.1.5	Représentation hiérarchique .....	34
9.1.2	<i>Construction du profil</i> .....	34
9.1.3	<i>Evolution du profil</i> .....	34
9.2	SELECTION DE L'INFORMATION .....	35
9.2.1	<i>Identification du profil</i> .....	35
9.2.2	<i>Exécution des requêtes</i> .....	35
9.2.3	<i>Présentation des résultats</i> .....	36
10.	LES SERVICES DE PERSONNALISATION .....	36
10.1	ENRICHISSEMENT DES REQUETES .....	36
10.2	LE FILTRAGE DES RESULTATS .....	36
10.3	LE RE-ORDONNANCEMENT DES ELEMENTS DU RESULTAT.....	37
10.4	LA RECOMMANDATION D'OBJET DU RESULTAT .....	37
11.	EXEMPLE DE SYSTEME DE GESTION DE PROFIL UTILISATEUR.....	37
12.	INTEGRATION DE L'USAGE DANS LA RECHERCHE D'INFORMATION.....	38
12.1	DEFINITION .....	38
12.2	WEB USAGE MINING (WUM) .....	38

12.3	ANALYSE DES USAGES.....	39
12.4	METHODES D'ANALYSE DES USAGES.....	39
12.4.1	<i>Analyse des traces logs</i> .....	39
12.4.2	<i>Analyse informatique des traces logs</i> .....	39
13.	AVANTAGES D'INTEGRATION DE L'USAGE.....	40
14.	CONCLUSION.....	40

### **Chapitre III: CONCEPTION**

1.	INTRODUCTION.....	41
2.	ARCHITECTURE DU SYSTEME.....	41
2.1	MODULE D'INDEXATION.....	43
2.2	MODULE D'ACQUISITION.....	43
2.2.1	<i>Acquisition explicite</i> .....	43
2.2.2	<i>Acquisition implicite</i> .....	43
2.3	MODULE DE RECHERCHE.....	43
2.4	MODULE D'ENRICHISSEMENT.....	43
3.	MODELISATION DU PROFIL UTILISATEUR.....	43
3.1	CHOIX D'UNE REPRESENTATION DU PROFIL.....	44
3.2	DESCRIPTION DES DIMENSIONS DE PROFIL UTILISATEUR.....	45
3.2.1	<i>Informations personnelles</i> .....	45
3.2.2	<i>Préférences</i> .....	45
3.2.3	<i>Centres d'intérêt</i> .....	46
3.2.4	<i>Usage</i> .....	46
3.2.5	<i>Construction de l'ontologie du domaine</i> .....	47
3.2.6	<i>Feedback (retour de pertinence)</i> .....	50
4.	FONCTIONNEMENT DU SYSTEME.....	50
4.1	MODULE D'INDEXATION.....	50
4.1.1	<i>Analyse lexicale</i> .....	51
4.1.2	<i>Elimination des mots vides</i> .....	51
4.1.3	<i>Lemmatisation (normalisation)</i> .....	51
4.1.4	<i>Pondération</i> .....	51
4.1.5	<i>Index</i> .....	52
4.2	MODULE D'ACQUISITION (GESTION DE PROFIL).....	52
4.2.1	<i>Acquisition explicite</i> .....	53
4.2.2	<i>Acquisition implicite</i> .....	53
4.2.2.1	<i>Processus d'observation</i> .....	54
4.3	MODULE DE RECHERCHE.....	56
4.4	MODULE D'ENRICHISSEMENT DE L'ONTOLOGIE.....	58
4.5	MODULE DE TRAITEMENT.....	58
4.5.1	<i>Traitement des requêtes</i> .....	58
5.	LES DIAGRAMMES UML REPRESENTATIFS DU SYSTEME.....	59
5.1	DIAGRAMME DES CLASSES.....	59
5.1.1	<i>Description des classes du système (Diagramme de classes)</i> .....	62
5.2	DIAGRAMME DES CAS D'UTILISATION.....	63
5.2.1	<i>Identification des cas d'utilisation</i> .....	63
5.2.2	<i>Description des cas d'utilisation</i> .....	63
5.2.3	<i>Diagramme de cas d'utilisation</i> .....	64
5.3	DIAGRAMMES DE SEQUENCES.....	64
5.3.1	<i>Diagramme de séquences du cas « s'inscrire »</i> .....	65
5.3.2	<i>Diagramme de séquences du cas « s'identifier »</i> .....	65
6.	CONCLUSION.....	66

### **Chapitre IV: MISE EN ŒUVRE ET TESTS**



1.	INTRODUCTION.....	67
2.	ENVIRONNEMENT DE DEVELOPPEMENT .....	67
2.1	JSP.....	67
2.2	APACHE TOMCAT 6.0.26 .....	67
2.3	NETBEANS 6.9.1.....	68
2.4	PROTEGE2000 .....	68
2.5	JENA .....	69
2.6	SGBD ORACLE .....	69
2.6.1	<i>Les caractéristiques d'Oracle</i> .....	70
3.	IMPLEMENTATION DU SYSTEME RC_PU_CU.....	71
3.1	CLASSES CORRESPONDANT AU MODULE "ACQUISITION DE PROFIL UTILISATEUR" .....	71
3.2	CLASSES CORRESPONDANT AU MODULE "LA RECHERCHE" .....	71
3.3	CLASSES CORRESPONDANT AU MODULE "ACQUISITION IMPLICITE DU PROFIL" .....	72
4.	INTERFACES DE L'APPLICATION.....	72
5.	EVALUATION DE PERTINENCE .....	80
6.	CONCLUSION .....	82
	<b>CONCLUSION GENERALE ET PERSPECTIVES</b> .....	83
	REFERENCES BIBLIOGRAPHIQUES.....	85
	ANNEXE A.....	88