

الأعلى

*République Algérienne Démocratique et Populaire*  
*Ministère de l'Enseignement supérieur et de la Recherche Scientifique*  
**Université des Sciences et de la Technologie Houari BOUMEDIENE**

**INSTITUT D'INFORMATIQUE**

Mémoire du Projet de fin d'études

Pour l'obtention du diplôme  
d'ingénieur d'état en **INFORMATIQUE**

Option : Software

SUJET :

**Conception et Réalisation d'un Système de Recherche  
d'Informations sur le Web basé sur la Reformulation de  
Requête**

**Thème proposé et encadré par :**

M<sup>me</sup> Z.ALIMAZIGHI (U.S.T.H.B)  
M<sup>me</sup> H.ALIANE (C.E.R.I.S.T)

**Étudié par :**

M<sup>elle</sup> BOUGHACHA RIME  
M<sup>er</sup> DJELLIOUT TOUFIK

Soutenu le : 31 Octobre 2001

**Devant le jury composé de :**

M<sup>me</sup> Mokhtari (*Présidente de jury*)  
M<sup>me</sup> Gherbi (*Membre de jury*)  
M<sup>me</sup> Akli (*Membre de jury*)

**PROMOTION :2000/2001**

## **Résumé :**

L'évolution des réseaux de transmission, en particulier l'Internet, nous emmène à manipuler quotidiennement un flot croissant d'informations multimédia. Mais cet océan de données serait bien inutile sans la capacité de les manipuler et d'y accéder rapidement et sélectivement.

Dans ce travail, nous proposons une approche pour la reformulation de la requête d'un utilisateur dans lequel nous faisons recours à la reformulation interactive. Cette approche doit impérativement contourner les problèmes d'ambiguïté et d'augmenter la précision des réponses.

Ainsi nous proposons dans ce mémoire une vue structurée d'un thème de recherche concernant la recherche d'informations sur le Web. L'objectif du travail est la conception et réalisation d'un système multi-agents fondé sur l'utilisation du profil d'utilisateur.

***Mots-clefs :*** Recherche d'informations, Web, reformulation de requête, système multi-agents, profil utilisateur, représentation de connaissances, algorithme d'apprentissage, client/serveur.

## Sommaire :

<i>Introduction générale</i> .....	02
<b><i>Chapitre I: Acquisition de connaissances</i></b> .....	<b>04</b>
I. Notions d'acquisition de connaissances.....	05
II. Les formalismes de représentation de la connaissance.....	06
II.1 La logique des propositions.....	06
II.2 Le calcul des prédicats.....	07
II.3 Les procédures.....	07
II.4 Les règles de production.....	08
II.4.1 La base des règles de production.....	08
II.4.2 Base des faits.....	08
II.4.3 Interpréteur des faits et des règles de production.....	09
II.5 Arbres de décisions.....	09
II.5.1 La forme d'une règle.....	09
II.5.2 L'analogie avec les arbres.....	09
II.5.3 L'algorithme ID3.....	10
II.6 Les réseaux sémantiques.....	10
II.6.1 Structure d'un réseau sémantique.....	10
II.6.2 Le filtrage.....	11
II.6.3 Représentation des concepts et héritage.....	12
II.6.4 Les inconvénients et les avantages des réseaux sémantiques.....	12
II.7 L'orienté objet.....	12
II.7.1 Fondements du modèle à objets.....	12
II.7.2 Les classes.....	13
II.7.2.1 Définition d'une classe.....	13
II.7.2.2 Relations entre classe.....	13
II.7.3 Les frames.....	14
II.7.4 Les acteurs.....	14
II.7.5 Les représentations hybrides.....	14
II.7.6 Intérêt du modèle à objet.....	14
<b><i>Chapitre II: Les Agents</i></b> .....	<b>15</b>
I. Intelligence Artificielle Distribuée.....	16
I.1 Les branches de l'IAD.....	16
II. Agents.....	17
II.1. Définition d'un agent.....	17
II.2. Caractéristique d'un agent.....	17
II.3 Modèles d'agents.....	18
II.3.1 Agents cognitifs.....	18
II.3.1.1 Structure interne.....	18
II.3.1.2 Fonctionnement.....	19
II.3.2 Agents réactifs.....	20
III. Système multi-agents.....	20
III.1 Définition de système multi-agents.....	20
III.2 Modèles de système multi-agents.....	20

III.2.1	Les systèmes multi-agents à contrôle centralisé .....	21
III.2.2	Systèmes multi-agents à contrôle distribué (Acteurs).....	21
III.2.2.1	Modèle d'acteurs de Hewitt.....	22
III.2.2.2	Modèle d'acteurs de Gul Agha.....	22
IV.	Société d'agents.....	23
IV.1	Coopération.....	23
IV.1.1	Coopération par partage de résultats.....	23
IV.1.2	Coopération par répartition ou partage de tâches.....	23
IV.1.2.1	Allocation centralisée.....	23
IV.1.2.2	Allocation distribuée des tâches.....	24
IV.1.2.2.1	Le réseau d'accointances.....	24
IV.1.2.2.2	Le réseau contractuel.....	24
IV.2	Résolutions de conflits.....	25
IV.2.1	Coordination.....	25
IV.2.2	Négociation.....	25
IV.2.2.1	Négociation centralisée.....	25
IV.2.2.2	Négociation distribuée.....	26
IV.3	Communication.....	26
IV.3.1	Protocole de communication.....	26
IV.3.2	Architectures de communication.....	27
IV.3.3	Les modes de communication.....	27
IV.3.3.1	Communication par partage d'informations.....	27
IV.3.3.2	Communication par envoi de messages.....	28
IV.4	Emergence.....	28
V.	Applications des SMA.....	28
VI.	Environnements de développement des SMA.....	29
 <b><u>Chapitre III : La Technologie Internet</u></b> .....		<b>30</b>
I.	Le réseau Internet.....	31
I.1	Présentation de l'Internet.....	31
I.2	Les services de l'Internet.....	31
I.2.1	Le courrier électronique (E-Mail).....	31
I.2.2	Le transfert des fichiers (FTP).....	31
I.2.3	La connexion à distance.....	31
I.2.4	Le service Wais.....	31
I.2.5	Le service Gopher.....	32
I.2.6	Le service Web.....	32
I.3	Les protocoles de l'Internet.....	32
I.3.1	Protocoles réseau.....	32
I.3.2	Protocole d'accès réseau.....	32
I.3.3	Protocole de transport.....	33
I.4	Avantages de TCP/IP.....	33
I.5	Le routage sur Internet.....	33
I.6	L'adressage sur Internet.....	34
I.6.1	L'adresse IP.....	34
I.6.2	Le nom du hôte.....	34
I.6.3	Les serveurs DNS(Domaine Name Server).....	35
II.	Le Web.....	35
II.1	Présentation du Web.....	35

II.2	Caractéristiques du Web.....	35
II.3	Les composants du Web.....	36
II.3.1	L'hypertexte.....	36
II.3.2	Les navigateurs Web.....	36
II.3.3	Le serveur Web.....	36
II.3.4	Les URL.....	36
II.3.5	Le protocole HTTP.....	36
II.3.6	Le langage HTML.....	36
II.4	Les applications Web.....	37
II.4.1	Les applications côté client.....	37
II.4.2	Les applications côté serveur.....	38
III.	Modèle Client / Serveur.....	38
III.1.	Définition.....	38
III.2.	Communication par messages.....	39
III.3.	Principe.....	39
III.4.	Mise en œuvre du client / serveur à l'aide des sockets.....	40
III.4.1	Socket.....	40
III.4.2	Notion de port.....	40
III.4.3	Enchaînement des opérations.....	40
III.4.3.1	Mode connecté.....	40
III.4.3.2	Mode non connecté.....	41
	<b><u>Chapitre IV : Recherche d'Information.....</u></b>	<b>43</b>
I.	Recherche d'information sur Internet.....	44
I.1	Les répertoires.....	44
I.2	Moteurs de recherche.....	45
I.3	Les méta-moteurs.....	45
I.4	Agents de recherche d'information.....	45
I.4.1	Les agents guides.....	46
I.4.2	Le systèmes de recommandation.....	46
II.	Les différents types de recherche.....	47
II.1	Recherches séquentielles ou indexées.....	47
II.2	Liens statiques ou virtuels.....	47
II.3	Recherche exactes ou approchées.....	47
II.4	Recherche par navigation ou requêtes.....	48
III	La reformulation de requête.....	48
III.1.	Les Typologies des reformulations de requête.....	48
III.1.1.	La reformulation manuelle.....	48
III.1.2.	La reformulation automatique.....	49
III.1.3.	La reformulation interactive.....	49
III.2.	Systèmes experts de reformulation de requête.....	49
III.3	Des études de stratégies de recherche.....	50
IV.	Indexation et modèle de représentation des documents.....	51
IV.1	L'indexation.....	51
IV.2	Représentation des documents.....	52
V.	Algorithmes d'apprentissage.....	54
V.1	L'algorithme k-Nearest Neighbor.....	54
V.2	Algorithme de classification bayésien naïf.....	55
VI.	L'extraction automatique de connaissances.....	56

<b>Chapitre V : Conception et réalisation.....</b>	<b>58</b>
I. Conception générale du système.....	59
I.1 Architecture générale du système.....	59
I.2 Description des composants du système.....	61
I.3 Fonctionnement du système.....	62
I.4 Mécanisme de communication Inter-Agent.....	62
II. Architecture du système de Ref-Requête et Recherche d'Information.....	63
II.1. Les Interfaces Interactives.....	64
II.1.1 Interface utilisateur.....	64
II.1.2 Interface Administrateur.....	64
II.2. La base de connaissances.....	64
II.2.1 Type de domaine de la connaissance.....	64
II.2.2 Les outils de gestion de la BC.....	66
II.3. Profil utilisateur.....	66
II.4. Base de données.....	67
II.5. Agent de reformulation de requête.....	67
II.5.1. La requête initiale.....	67
II.5.2. Le processus de reformulation de la requête.....	68
II.5.2.1 Principes de base de la reformulation.....	68
II.5.2.2 Extension de la requête.....	68
II.5.2.3 Le processus de reformulation.....	69
II.6. Agent collecteur de documents.....	69
II.6.1 Accès aux données via Les scripts CGI .....	70
III. L'interaction avec le système.....	70
III.1 L'application serveur.....	71
III.2 Application client.....	71
III.4 Cas de plusieurs utilisateur.....	72
<b>Chapitre VI : Description du logiciel.....</b>	<b>73</b>
I. Outils de développement.....	74
II. Description du logiciel .....	75
<b>Conclusion générale.....</b>	<b>85</b>
<b>Annexes.....</b>	<b>86</b>
Annexe 1 : Langage JAVA.....	87
Annexe 2 : Langage Perl.....	91
Annexe 3 : Les threads.....	93
<b>Glossaire.....</b>	<b>95</b>
<b>Bibliographie.....</b>	<b>97</b>