

République Algérienne Démocratique et Populaire
Ministère de L'Enseignement Supérieur et de la Recherche Scientifique

Université des Sciences et de la Technologie Houari BOUMEDIENE



Faculté d'Electronique et Informatique
Département Informatique

Mémoire du Projet de fin d'études

Pour l'obtention du diplôme
d'ingénieur d'état en Informatique

SUJET :

Systeme d'information fédéré pour la recherche documentaire multibase

Proposé et dirigé par :
Mr. MOSTEFAI Mohamed Amine

Etudié par :
BOUKHEBOUZE Mohamed
FERRAG Hamza

Co-promoteur :
Mr. BATATA Sofiane

Soutenu le : 09/10/04

Devant le jury composé de :

Mme. MAMERI	(Présidente)
Mr. BOUKRA	(Examineur)
Melle. BENCHENNAF	(Examinatrice)

PROMOTION : 90/2004

Sommaire

Introduction	1
---------------------------	---

Chapitre 1 : Recherche d'information documentaire sur Internet

1. Introduction.....	4
2. Problématique de la recherche d'information.....	4
3. Recherche d'information sur le Web.....	5
3.1 Les Annuaire.....	5
3.2 Les moteurs de recherche.....	6
3.3 Comparaison entre les moteurs de recherche et les annuaires.....	7
3.4 Le Web statique et le Web dynamique.....	8
4. Recherche dans une base de données.....	9
5. Recherche documentaire.....	11
5.1 Organisation du réseau informatique au sein d'une bibliothèque.....	12
5.2 Le catalogage.....	13
5.3 Les catalogues en ligne.....	14
5.4 Structure du format UNIMARC.....	15
5.5 Types de recherche documentaire.....	16
6. Conclusion.....	18

Chapitre 2 : Les systèmes d'information fédérés (SIF)

1. Introduction.....	21
2. Classification des systèmes d'information.....	21
2.1 Critères de classification des systèmes d'information.....	21
2.1.1 L'hétérogénéité.....	21
2.1.2 L'autonomie.....	23
2.1.3 La distribution.....	24
2.2 Types de systèmes d'information.....	24
3. Architecture générale d'un système d'information fédéré.....	24
4. Méta-données pour les systèmes d'information fédérés.....	26
5. Quelques critères de classification pour les SIFs.....	27
5.1 Types de composants.....	27
5.2 Fédération « serrée » et fédération « lâche ».....	27
5.3 Transparence.....	28
5.4 L'approche ascendante et l'approche descendante.....	29
5.5 Intégration virtuelle et intégration matérialisée.....	29
6. Types de systèmes d'information fédérés.....	30
6.1 Systèmes d'information faiblement couplés.....	30
6.2 Systèmes d'information fortement couplés.....	31
6.2.1 Bases de données fédérées.....	31
6.2.2 Systèmes d'information basés sur le médiateur.....	33
7. Synthèse.....	36
8. Conclusion.....	36

Chapitre 3 :**Partie 1 : XML et les bases de données**

1. Introduction.....	39
2. DTD.....	40
3. XML-Schema.....	41
4. XML et l'interopérabilité entre les bases de données.....	42
4.1 Données et Documents.....	43
4.1.1 Les contenus orientés données.....	43
4.1.2 Les contenus orientés documents.....	43
4.2 La correspondance entre les schémas de documents et les schémas de bases de données.....	45
4.2.1 La correspondance basée sur des tables.....	45
4.2.2 La correspondance basée sur un modèle objet relationnel.....	46

Partie 2 : Les technologies distribuées et services Web

1. Introduction.....	48
2. L'architecture CORBA.....	48
2.1 Présentation et historique.....	48
2.2 Architecture de CORBA.....	49
2.2.1 Object Request Broker (ORB).....	49
2.2.2 Les interfaces d'invocation.....	49
2.2.3 Le référentiel d'interface.....	50
2.2.4 Les souches de communication (<i>stubs</i> et <i>skeletons</i>).....	50
2.2.5 L'adaptateur d'objets.....	50
3. Les technologies COM/DCOM.....	51
3.1 Présentation et historique.....	51
3.2 L'architecture de COM.....	51
3.2.1 Interface COM.....	52
3.2.2 Les serveurs COM.....	52
3.2.3 Les clients COM.....	54
4. La technologie Java RMI.....	55
4.1 Introduction.....	55
4.2 Principe de fonctionnement.....	55
4.2.1 Les interfaces.....	55
4.2.2 Le modèle en couche.....	55
5. SOAP et les Services Web.....	57
5.1 Présentation et fonctionnement.....	57
5.2 Présentation et fonctionnement des Services Web.....	57
5.3 Protocoles utilisés par SOAP.....	58
5.4 Le langage WSDL.....	59
5.5 Le protocole UDDI.....	61
6. La technologie .NET.....	61
6.1 L'environnement d'exécution .NET.....	61
6.2 Présentation de .NET Remoting Framework.....	63
6.2.1 Les formateurs.....	63
6.2.2 Les types de canaux.....	64

6.3 Activation.....	64
7. Synthèse.....	65
8. Conclusion.....	66

Chapitre 4 : Conception

1. Introduction.....	69
2. Architecture générale du système.....	69
2.1 Une architecture multi-niveaux.....	69
2.2 Modèle de données.....	70
2.3 Prototype du système.....	73
2.3.1 Description détaillée du médiateur.....	74
2.3.2 Description détaillée de l'adaptateur.....	76
2.3.3 Description détaillée de l'application cliente.....	82
2.3.4 La communication entre les différents modules.....	84
3. Conclusion.....	86

Chapitre 5 : Réalisation

1. Introduction.....	88
2. Environnement de développement.....	88
3. Implémentation des bases de données.....	89
4. Implémentation du médiateur.....	90
5. Implémentation de l'adaptateur.....	94
5.1 L'adaptateur pour bases de données relationnelles.....	97
5.1 L'adaptateur pour fichiers <i>.OUV</i> de <i>SYNGEB</i>	101
6. Implémentation de l'application cliente.....	102
7. La communication entre les différents modules du système.....	105
8. Le temps de réponse.....	108
9. L'extension de l'architecture.....	108

Conclusion et perspectives.....	110
--	------------

ANNEXE A : Le format UNIMARC.....	112
--	------------

ANNEXE B : Histoire du catalogue.....	120
--	------------

ANNEXE C : Présentation de XML.....	123
--	------------

Références bibliographiques.....	138
---	------------