

République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université de la science 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 : Bases de données / technologie du Web

Thème

Interrogation d'une base de documents XML hétérogène
en utilisant le langage XQuery

Proposé et encadré par : Madame Bessai F.Z (CERIST).

Etudié par : Bey Sihem .

Sayah Mohamed Abd-El-Aziz.

Presidente du jury : Mokhtari Aicha

Membres du jury : Bouagar Saliha

Kechid Samir



01 Juillet 2006

Table des matières

Introduction	2
Chapitre I – Le langage XML	
I.1 – Introduction	5
I.2- Quels sont les composants de XML ?	6
I.3- Parseur XML	6
I.4- Conception d'un document XML bien formé	7
I.4.1- Créer des éléments XML	7
I.4.2- Les attributs	8
I.4.3- Les commentaires	9
I.4.4- Déclaration XML	9
I.4.5- Les instructions de traitement	10
I.4.6- Appels d'entités et sections CDATA	10
I.5- Les Définitions de type de document (ou DTD)	11
I.6- Les espaces de noms	12
I.6.1- Déclaration des espaces de noms	12
I.6.2- Utilisation des espaces de noms	12
I.7- Les feuilles de style	13
I.8- XSLT	13
I.9- XPATH	13
I.10- Liens XML	13
I.11- DOM	14
I.12- SAX	14
I.13- Les schémas XML	15
I.14- Conclusion	16
Chapitre II – XQuery : Langage d'interrogation de documents XML	
II.1 – Introduction	18
II.2- Les expressions XQuery	18
II.2.1- Les expressions de définition de chemin d'accès	18
II.2.2- Constructeurs d'éléments	20
II.2.3- Les commentaires	21
II.2.4- Les expressions FLWOR	21
II.2.5- Les opérateurs dans les expressions	25
II.2.6- Les expressions conditionnelles (IF THEN ELSE)	25
II.2.7- Les quantificateurs	26
II.2.8- Le filtrage	26

II.2.9- Types de données	27
II.2.10- La déclaration des types complexes	27
II.2.11- Les opérations sur les types	28
II.2.12- Utilisation des parenthèses	29
II.2.13- Priorité des opérateurs	30
II.3- Prologues et Modules	30
II.3.1- Déclaration de la version de XQuery	30
II.3.2- Déclaration de module	31
II.3.3- Importation de Modules	31
II.3.4- Importation de schémas	31
II.3.5- Déclaration de variables	32
II.3.6- Déclaration de fonctions	33
II.3.7- Déclaration d'espaces de noms	35
II.4- Quelques opérations	36
II.5- Conclusion	39
Chapitre III – Conception de l'application	
III.1- Introduction	41
III.2- Le langage de modélisation unifiée (UML)	41
III.2.1- Diagramme de cas d'utilisation	42
III.2.1.1- Utilisateurs – Acteurs	42
III.2.1.2- Les cas d'utilisation	43
III.2.1.3- Les relations entre cas d'utilisation	43
III.2.2- Diagramme d'états/transition	44
III.2.2.1-états, transition, événement et action	45
III.3- Architecture du prototype d'interrogation d'une base de documents XML hétérogène	45
III.3.1-Le module « Importation »	46
III.3.1.1- Importation d'un schéma XML	47
III.3.1.2- Importation d'un document XML	47
III.3.2 – Représentation de la base de documents	49
III.3.3- Le module « Sélection d'une classe de documents »	50
III.3.4 : Le module « Interrogation »	50
III.3.5- le module«Enregistrer résultats»	53
III.3.6- le module«Afficher résultats»	53
III.4-Conclusion	54
Chapitre IV – Implémentation de l'application	
IV.1- Introduction	56
IV.2- Langage de programmation utilisé	57

IV.3- Contraintes logicielles et matérielles	57
IV.3.1- Contraintes logicielles	57
IV.3.2- Contraintes matérielles	57
IV.4- L'implémentation 'AltovaXML 2006'	57
IV.5 - Description de notre système d'interrogation d'une base de documents XML hétérogènes	58
IV.5.1- L'interface principale de notre système d'interrogation	58
IV.5.2- Menu 'Importation'	59
IV.5.2.1- Menu 'Importer un document'	59
IV.5.2.2- Menu 'Importer un schéma XML'	62
IV.5.3- Menu 'Interrogation'	62
IV.5.3.1- Interrogation par contenus	62
IV.5.3.2- Interrogation par contraintes structurelles simples	67
IV.5.3.3- Interrogation par contraintes structurelles avancées	70
IV.6- Conclusion	73
Conclusion et perspectives	74
Glossaire	76
Références bibliographiques	80
Annexe A – Les DTD	86
Annexe B – XML Schéma	102
Annexe C – XPath	118
Annexe D – Les feuilles de styles	126

Résumé

Le développement du document électronique et du Web ont vu émerger puis s'imposer des formats de données structurés, tels que le SGML (Standard Generalized Markup Language) et le XML (eXtensible Markup Language), permettant de représenter l'information sous une forme plus riche que le simple contenu et adaptée à des besoins spécifiques. Ces nouveaux formats permettent de représenter conjointement l'information textuelle et l'information de structure d'un document.

Pour pouvoir interroger des documents XML, il faut d'abord connaître leurs structures, qui sont souvent hétérogènes mais qui peuvent être définies par des Schémas XML. La réalisation d'un système d'interrogation qui permet la recherche des informations sur une base de documents hétérogènes s'avère nécessaire. Ce système doit pouvoir générer une interface de recherche pour n'importe quel schéma XML et pouvoir formuler et exécuter des requêtes XQuery (XML query) selon les informations saisies, ce qui permettra aux utilisateurs d'interagir avec la base de documents XML hétérogène

Mots Clés : Données structurées, XML, XML Schéma, XPath, XQuery, Base de Données.