

*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 : Bases de Données / Web

### Thème

**Stockage et Interrogation de documents XML  
avec le langage XQuery**

proposé par : Mme Bessaï F.Z (CERIST)

Etudié par :

SLIMANI Farid

Encadré par : Mme Bessaï F.Z (CERIST)

BELMANSOUR A. Tidjani

Soutenu le : 02 Octobre 2004

Devant le jury composé de :

Mlle Bensaou

Président,

Mr Boukra

Examinateur,

Mr Boumaza

Examinateur,

PROMOTION : 2003 / 2004

N° de projet : 55 / 2004

# Table des matières

<b>Introduction</b>	1
<b>Chapitre I – Le langage XML</b>	
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
<b>Chapitre II – XQuery : langage d'interrogation de documents XML</b>	
II.1- Introduction	18
II.2- Les expressions XQuery	19
II.2.1- Les expressions de définition de chemin d'accès	19
II.2.2- Constructeurs d'éléments	21
II.2.3- Les commentaires	22
II.2.4- Les expressions FLWOR	22
II.2.5- Les opérateurs dans les expressions	26
II.2.6- Les expressions conditionnelles (IF THEN ELSE)	26
II.2.7- Les quantifieurs	26
II.2.8- Le filtrage	27
II.2.9- Types de données	27
II.2.10- La déclaration des types complexes	28
II.2.11- Les opérations sur les types	29
II.2.12- Utilisation des parenthèses	30
II.2.13- Priorité des opérateurs	30
II.3- Prologues et Modules	31
II.3.1- Déclaration de la version de XQuery	31
II.3.2- Déclaration de module	31
II.3.3- Importation de Modules	32
II.3.4- Importation de schémas	32

II.3.5- Déclaration de variables	33
II.3.6- Déclaration de fonctions	34
II.3.7- Déclaration d'espaces de noms	36
II.4- Quelques opérations	37
II.5- Conclusion	40
<b>Chapitre III – Conception du prototype « XSearcher »</b>	
III.1- Introduction	43
III.2- Le langage de modélisation unifié (UML)	43
III.3- Architecture de « XSearcher »	45
III.3.1- Le module d'importation	46
III.3.1.1- Pourquoi nous restreindre à une base de documents de livres et ne pas concevoir un système générale d'interrogation de documents XML?	49
III.3.1.2- Structure d'un document XML représentant un livre	49
III.3.1.3- Schéma XML associé à un livre	50
III.3.1.4- Schéma de la base de documents	51
III.3.2- Le module d'interrogation	52
III.3.3- Le module d'affichage des résultats	55
III.3.4- Module d'enregistrement	57
III.4-Conclusion	58
<b>Chapitre IV – Implémentation du prototype « XSearcher »</b>	
IV.1- Introduction	60
IV.2- Implémentation XQuery utilisée	60
IV.3- Langage de programmation utilisé	62
IV.4- Contraintes logicielles et matérielles	62
IV.4.1- Contraintes logicielles	62
IV.4.2- Contraintes matérielles	62
IV.5- Le prototype « XSearcher »	62
IV.5.1- L'interface principale de « XSearcher »	63
IV.5.2- Importation de documents XML	63
IV.5.2.1- Importer un document XML	63
IV.5.2.2- Importer un dossier contenant des documents XML	65
IV.5.3- Interrogation des documents XML	67
IV.5.3.1- Interrogation par contenus	67
IV.5.3.2- Interrogation par contraintes structurelles simples	72
IV.5.3.3- Interrogation par contraintes structurelles avancées	75
IV.6- Conclusion	79
<b>Conclusion</b>	82
<b>Glossaire</b>	85
<b>Références bibliographiques</b>	88
<b>Annexe A – Les DTD</b>	91

<b>Annexe B – XML Schéma</b>	107
<b>Annexe C – XPath</b>	123
<b>Annexe D – Les feuilles de style</b>	131

# Résumé

Le développement du document électronique et du Web ont vu émerger puis s'imposer des formats de données structurées 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.

Ce qu'on appelle la famille XML, se sont les standards de base qui s'appuient sur XML : XML Schema pour les modèles de documents, XPath pour le cheminement dans les arbres XML, XSL pour les feuilles de styles, XQuery pour l'interrogation de documents, DOM et SAX pour la programmation, SOAP pour les services distribués, etc.

Les standards d'adressage (XPath), de transformation (XSLT) et d'interrogation (XQuery) sont de plus en plus utilisés et il est vrai que le développement d'XML passe par le développement d'outils (tels qu'un système de gestion de base de données XML native) et de composants qui permettent de travailler facilement avec XML. Ces outils doivent s'appuyer sur ces standards de base.

Nous allons particulièrement nous intéresser aux possibilités d'interrogation offertes par le langage XQuery pour les documents structurés en XML. Par conséquent, notre travail consistera en la conception et la mise en œuvre d'une base de documents XML ainsi qu'une interface d'interrogation, basée sur le langage XQuery, qui permettra aux utilisateurs d'interagir avec cette base.

**Mots Clés :** Données structurées, XML, XML Schema, XPath, XQuery, Base de Données, stockage, interrogation par contenus, interrogation par contraintes structurelles.