



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'électronique et d'informatique

Département d'informatique

Mémoire du projet de fin d'études pour l'obtention

d'ingénieur d'état en informatique

Option : Technologie du Web

Thème :

Recherche d'information dans une base de documents XML

basé sur le langage XQuery Full Text

Propose et encadre par : Mme Bessai F.Z (CERIST)

Etudie par :

Redam Nacim

Berezg Tarek

Sommaire

Sommaire

Introduction générale	1
Chapitre 01 : Le langage XML	
1. Introduction	5
2. Généralités sur le langage XML	5
2.1. Qu'est-ce que le XML?	5
2.2. Quelles sont les principales utilisations de XML ?	5
2.3. Qui décide pour le XML et les langages associés ?	6
2.4. Pourquoi XML?	6
2.5. Quels sont les composants de XML ?	7
2.6. Document XML	9
2.7. Qu'est ce qu'un document XML bien formé ?	10
2.8. Qu'est ce qu'un document XML valide ?	11
2.9. Les parseurs	11
2.10. Les Définitions de type de document (ou DTD)	12
2.11. XPATH	12
2.12. Les feuilles de style	13
2.13. XML Schéma	14
3. Conclusion	15
Chapitre 02 : Le langage XQuery Full Text	
1. Introduction	17
2. Définition	17
3. Les expressions XQuery Full Text	18
3.1. Les expressions de définition de chemins d'accès	18
3.2. Constructeurs d'éléments	20
4. Les commentaires	22
5. Expression <i>FTContains</i>	22
5.1. Full Text Selection (FTSelection)	22
5.1. a. La sélection primaire du Full Text	23
5.1. b. FTSelection élémentaire (FTWords)	24
5.1. c. Les options du Match	25

5.1. d. Opérateurs logique du Full Text	31
5.1. e. Les filtres positionnels (Positional Filters)	33
5.1. f. Sélection de la Cardinalité (Cardinality Selection)	36
5.2. Option Ignorer (Ignore Option)	37
6. Les expressions FLOWR	38
7. Les opérateurs dans les expressions	43
8. Les expressions conditionnelles (IF THEN ELSE)	44
9. Les quantificateurs	44
10. Le filtrage	44
11. Types de données	45
12. Priorité des opérateurs	46
13. Prologue et module	47
13.1. Extension du contexte statique dans la version Full Text	47
13.2. Déclaration de la version de XQuery	47
13.3. Déclaration de module	47
13.4. Importation de Modules	48
13.5. Importation de schémas	48
13.6. Déclaration de variables	49
13.7. Déclaration de fonctions	50
13.8. Déclaration d'espaces de noms	52
14. Conclusion	53

Chapitre 03 : Conception d'un système d'interrogation de

documents XML basé sur le langage XQuery Full-Text

1. Introduction	55
2. Le langage de modélisation unifié UML	55
2.1. Diagramme de cas d'utilisation (use case)	57
2.1.1. Acteur	57
2.1.2. Le cas d'utilisation	58
2.1.3. La relation	58
2.2. Diagramme d'états/transition	59
2.2.1. Etat	59

2.2.2. Evénements et transitions	60
2.2.3. Les traitements	60
2.2.4. Les états prédéfinis	60
2.3. Diagramme de classes	61
3. Architecture du prototype d'interrogation d'une base de documents XML	
basé sur le langage XQuery Full Text	61
3.1. Le module « Importation »	64
3.1.1. Importation d'un schéma XML	64
3.1.2. Importation d'un document XML	64
3.2. Le module « Interrogation »	66
3.2.1. Recherche par contenu	66
3.2.2. Recherche structurelle simple	66
3.2.3. Recherche structurelle avancée	67
3.3. Options du langage XQuery Full Text	69
3.3.1. La sélection de la méthode de recherche des mots	69
3.3.2. La sélection du contenu	69
3.3.3. La sélection de la fenêtre des mots (FTWindow)	69
3.3.4. La sélection des mots vides (stop words)	69
3.3.5. La définition de l'ordre des mots (FTOrder)	70
3.3.6. La définition des occurrences des mots	70
3.3.7. La définition de la distance entre les mots (FTDistance)	70
3.3.8. L'emplacement des mots	70
3.3.9. La sélection du domaine des mots	70
3.4. Le module«index »	71
3.5. Le module«score»	72
3.6. Le module « Enregistrer résultats »	72
3.7. Le module « Afficher résultats »	73
4. Diagramme de classe modelisant notre systeme d'interrogation de documents XML	
Utilisant le langage XQuery Full Text	73
5. Conclusion	75

Chapitre 04 : implémentation du système d'interrogation

1. Introduction	77
2. Langage de programmation utilisé	77
3. Contraintes matérielles et logicielles	78
3.1. Contraintes logicielles	78
3.2. Contraintes matérielles	78
4. L'implémentation 'AltovaXML 2008'	78
5. Description de notre système d'interrogation	79
5.1. Architecture du prototype d'interrogation d'une base de documents XML basé sur le langage XQuery Full Text	79
5.2. Interface principale du système	82
5.3. Le menu « importation »	82
5.3.1. Le menu « importation de schémas »	82
5.3.2. Le menu « importation de documents »	85
5.4. Le menu « interrogation »	86
5.4.1. Le menu « interrogation par mots clés »	87
5.4.2. Le menu « Interrogation structurelle simple »	89
5.4.3. Le menu « Interrogation structurelle avancée »	92
6. Conclusion	96
Conclusion générale	98
Références bibliographiques	100
Glossaire	104
Annexe A : les documents XML	108
Annexe B : les DTD	119
Annexe C : les feuilles de styles	132
Annexe D : les schémas XML	158
Annexe E : XPath	173