



N° d'ordre : .....

**UNIVERSITE DE M'SILA**  
**FACULTE DES MATHÉMATIQUES ET DE L'INFORMATIQUE**  
**Département d'Informatique**

**MEMOIRE**

**Présenté pour l'obtention du diplôme de Master**

**Domaine : Mathématiques et Informatique**

**Filière : Informatique**

**Spécialité : Systèmes d'informations avancées**

**Par :**

**Fatima BOUABDALLAH**

**SUJET**

**Conception et réalisation d'un système  
Questions/Réponses en langage naturel pour  
l'exploitation d'une base de données  
géographique**

**Soutenu publiquement le : 28/06/2012 devant le jury composé de :**

**Mr. A. MOUSSAOI**  
**Mr. R. MAHDJOUBI**  
**Mr. B. BRAHIMI**

**Université de M'sila**  
**Université de M'sila**  
**Université de M'sila**

**Président**  
**Rapporteur**  
**Examineur**

**Promotion : 2011 / 2012**

# SOMMAIRE

---

<b>Introduction générale:</b> .....	1
<b>Chapitre 01 : Généralités sur les Systèmes Questions-Réponses</b>	
1.1 Introduction .....	3
1.2 Comparaison entre moteur de recherche et SQR : .....	3
1.3 Historique : .....	4
1.4 Architecture d'un système de question-réponse : .....	5
1.4.1 Analyse des questions : .....	6
1.4.2 Recherche des documents : .....	8
1.4.3 Analyse des documents candidats : .....	8
1.4.4 Extraction des réponses : .....	8
1.5 Présentation de quelques systèmes de question-réponse : .....	9
1.5.1 Le système QALC : .....	9
1.5.2 Le système PIQUANT : .....	9
1.5.3 Le système WEBCOOP : .....	10
1.5.4 Le système QRISTAL : .....	10
1.6 Conclusion : .....	10
<b>Chapitre 02 : Etude conceptuelle et fonctionnelle</b>	
2.1 Introduction : .....	11
2.2 Présentation du langage UML (Unified Modeling Language) : .....	11
2.3 L'étude fonctionnelle : .....	12
2.2.2 Diagramme de cas d'utilisation : .....	13
2.2.2 Diagramme de séquence : .....	13
2.3.3 Diagramme états-transitions : .....	14
2.3.4 Diagramme de classes : .....	15
2.4 Conclusion : .....	19
<b>Chapitre 03 : Architecture et fonctionnement</b>	
3.1 Introduction : .....	20
3.2 Grammaire des questions : .....	20
3.3 La base de données : .....	21
3.4 Système QUAGéo : .....	22
3.4.1 L'analyse de la question : .....	23

3.4.2	La recherche et l'extraction de la réponse :.....	25
3.5	Conclusion :.....	26
<b>Chapitre 04 : Réalisation du système « QUAGéo »</b>		
4.1	Introduction : .....	27
4.2	Les outils de développement : .....	27
4.2.1	Microsoft.Net :.....	27
4.2.2	Visual studio :.....	27
4.2.3	SQL Server 2008 : .....	28
4.2.4	C# .Net :.....	28
4.2.5	L'outil Irony de .Net:.....	28
4.3	Description d'interface du système « QUAGéo » : .....	28
4.4	Présentation de quelques interfaces :.....	29
<b>Conclusion générale</b> .....		33
<b>Bibliographie</b> .....		34
<b>Annexe A</b> .....		35
<b>Annexe B</b> .....		36

## *Introduction générale*

---

La quantité de documents électroniques mise à disposition, notamment grâce aux réseaux informatiques, a largement modifié la notion de recherche d'information. Les utilisateurs ont en effet un accès de plus en plus direct à l'information. Cependant, pour accéder plus facilement à une information pertinente, des systèmes de recherche d'information se révèlent incontournables. Jusqu'aujourd'hui le marché de l'informatique essaie de répondre à cette problématique en développant des outils spécifiques tels que les moteurs de recherche, les systèmes de Question/Réponse, les systèmes d'extraction d'informations, les analyseurs morphologiques et syntaxiques... etc.

Bien que les moteurs de recherche constituent une solution efficace pour trouver des documents correspondant à une requête utilisateur, ils s'avèrent moins performants concernant la recherche d'une donnée précise. De ce fait, il est primordial de faire appel à des systèmes plus élaborés capables de retourner une information fiable à un besoin d'information précis. C'est l'ambition des systèmes de questions-réponses.

Les systèmes de Question/Réponse (Q/R) sont une extension des systèmes de recherche d'information allant dans ce sens. Ce type de systèmes permet à un utilisateur de poser une question en langage naturel et de recevoir une réponse adéquate et précise dans le même langage. Ces systèmes peuvent être appliqués à différents domaines desquels celui de la géographie.

Le domaine géographique dispose aujourd'hui d'une grande quantité de documents électroniques et de multiples ressources linguistiques et terminologiques, ce qui nécessite l'utilisation de moyens très efficaces pour les exploiter, comme le SIG (Système d'information géographique), l'atlas géographique mondial, les dictionnaires géographiques...etc. Ce vaste domaine est caractérisé par la richesse et la complexité de son vocabulaire spécialisé.

Dans ce projet, notre objectif est de réaliser un système de question-réponse pour interroger une base de données géographique afin d'obtenir une réponse précise à une question formulée par l'utilisateur

Pour aboutir à ce but, nous avons procédé à une étude détaillée exposée dans ce mémoire qui s'articule autour de quatre chapitres suivants :

- ✓ **Le premier chapitre** : donne une vue générale sur les systèmes de question-réponse en présentant quelques-uns d'entre eux.
- ✓ **Le deuxième chapitre** : présente l'étude fonctionnelle de notre système et décrit son étude conceptuelle à travers le langage de modélisation UML.
- ✓ **Le troisième chapitre** : a pour but de présenter les différentes étapes de développement du système « QUAGéo ».
- ✓ **Le quatrième chapitre** : présente les outils de développement de notre système, ainsi que la présentation de quelques-unes de ses interfaces.

## *Conclusion générale :*

---

Dans ce projet, nous nous sommes fixés au départ, la réalisation d'un système qui permettrait à quiconque de poser des questions en langage naturel pour obtenir en contrepartie, une réponse, également en langage naturel et ceci sans que l'utilisateur ne soit obligé d'apprendre un langage de requêtes spécifique. Finalement, nous pouvons dire que cet objectif a été plus ou moins atteint malgré quelques manques que, nous espérons, ne nuisent pas à la qualité du résultat et qu'ils puissent être facilement récupérables. Ces lacunes sont principalement dus au manque de travaux antérieurs sur le domaine, à la nouveauté des outils utilisés ainsi qu'à la courte période consacrée à cette étude.

Enfin, nous tenons à préciser qu'au cours de la réalisation de ce projet nous avons appris beaucoup de nouvelles connaissances que ce soit dans le domaine de la géographie ou celui des systèmes questions/réponses avec tout ce qui s'y rattache.

En résumé, nous pouvons dire que ça été une agréable expérience qui nous a aidé à découvrir notre potentiel et mettre sur le terrain tout ce qu'on a appris durant notre période de formation.

Toutefois, plusieurs perspectives et améliorations peuvent être envisagées afin de contribuer à la perfection de cette application. Nous citons à titre d'exemple :

- Utiliser d'autres plateformes.
- Elargir la base de données pour d'autres branches de la géographie.
- Introduire de nouvelles langues pour obtenir un système multi-langues.
- Introduire la notion d'ontologie géographique pour élargir l'éventail des concepts qui vont de pair avec d'autres.

## *Bibliographie & webographie*

---

### **Bibliographie :**

- [BEN04] Benamara F. Cooperative question answering in restricted domains: the Webcoop experiment. In: ACL-Portability of systems, Barcelona, MIT Press, 2004.
- [BEL06] Luc Bélanger : Architecture question-réponse pour l'automatisation des services d'information. Thèse présentée à la Faculté des études supérieures en vue de l'obtention du grade de Philosophiæ Doctor (Ph. D.) en informatique. Université de Montréal, 2006.
- [CHU02] Chu-Carroll J. & Prager J. & Welty C. & Czuba K. & Ferrucci D: A Multi-Strategy and Multi-Source approach to question answering. In: Proceedings of the 11<sup>th</sup> Text Retrieval Conference (TREC-11), 2002.
- [EMB08] Mehdi EMBAREK : Un système de question-réponse dans le domaine médical- le système Esculape. Thèse pour obtenir le grade de Docteur de l'Université Paris-Est, 2008.
- [FER00] Ferret O. & Grau B. & Hurault-Plantet M. & Illouz G. & Jacquemin C. & Masson N. & Lecuyer P.: QALC-the question answering system of LIMSI-CNRS. In: Technical report: LIMSI-CNRS TREC 9 evaluation, 2000.
- [LAU05] Laurent D. & Séguéla P.: Qristal, système de Question-Réponse. In Traitement Automatique des Langues Naturelles (TALN 2005), Dourdan, 2005.
- [ROQ08] PascalRoques: Les cahiers du programmeur, UML 2 Modéliser une application web. Eyrolles, 2008.
- [VIL11/12] F.-Y. Villemin : Introduction à UML 2.0. CNAM, MAI NFE103 ,2011-2012

### **Webographie :**

- [Web-SQLpro] SQL server. [sqlpro.developpez.com](http://sqlpro.developpez.com).
- [Wikipédia] les système questions-réponses. [fr.wikipedia.org](http://fr.wikipedia.org)

## ملخص :

هدف هاته المذكرة هو تطوير نظام سؤال-جواب « QUAGéo »، يسمح للمستخدم بطرح سؤال في الجغرافيا، ومن ثمة يعمل على ارجاع الجواب المطلوب بدقة. ولقد اخترنا لغة ال UML من اجل انشاء تصور لبنية هذا النظام، Microsoft SQL Server لانشاء قاعدة البيانات، Irony من اجل كتابة القواعد المتبعة في تحليل الأسئلة المطروحة و أخيرا .NET et Visual studio 2010 plate-forme لإنجاز هذا النظام.

**الكلمات المفتاح :** نظام سؤال-جواب، الجغرافيا، لغة طبيعية، قواعد تحليل الأسئلة، التركيب، UML، Microsoft SQL Server، Irony، Visual studio 2010، C#.

---

## Résumé :

Ce mémoire a pour objectif le développement d'un système question/réponse nommé « QUAGéo » qui permet à l'utilisateur de poser une question dans le domaine géographique en langage naturel afin d'obtenir une réponse précise. Nous avons choisi la méthode UML pour modéliser la structure et le comportement de notre système, Microsoft SQL Server pour développer la base de données, Irony pour développer la grammaire des questions et enfin, la plate-forme .NET et Visual studio 2010 pour la réalisation.

**Mots clés :** systèmes question-réponse (SQR), géographie, langage naturel (LN), UML (Unified Modeling Language), Microsoft SQL Server, Visual studio 2010, grammaire, Irony, syntaxe, C#.

---

## Summary:

This thesis has for objective the development of a question answering system, that we called QUAGéo, which allows the user to ask a question in the field of geography, in natural language, in order to get a precise response. We chose UML to model the structure and behavior of our system, Microsoft SQL Server to develop the database, Irony to develop the grammar of questions, and finally the .NET framework and Visual Studio 2010 to implement our system.

**Keywords:** Question-Answer Systems, Geography, Natural Language, UML (Unified Modeling Language), Microsoft SQL Server, Visual studio 2010, Grammar, Irony, Syntax, C#.