

UNIVERSITE DE CONSTANTINE

INSTITUT D'INFORMATIQUE

**THESE DE MAGISTER**

**Spécialité : INFORMATIQUE**

**Option : SYSTEME INFORMATIQUE**

Présentée par :

**Mohamed el hadi BENELHADJ**

THEME :

**LE TRAITEMENT DE TEXTE DANS LES  
BASES DE DONNEES RELATIONNELLES**

**( DONNEES BIBLIOGRAPHIQUES )**

Soutenu le 21 Janvier 1992

Devant le Jury :

<b>C. B. BEN-YELLES</b>	- Professeur	- USTHB	- Président
<b>J. E. JEACOCKE</b>	- Docteur	- Glasgow	- Rapporteur
<b>M. BETTAZ</b>	- Professeur	- Constantine	- Examineur
<b>Z. ALI-MAZIGHI</b>	- Docteur	- USTHB	- Examinatrice
<b>M. T. LASKRI</b>	- Docteur	- Annaba	- Examineur
<b>Z SAHNOUN</b>	- Docteur	- Constantine	- Examineur

## SOMMAIRE

<b>I. INTRODUCTION.....</b>	<b>1</b>
1.1. L'importance du texte.....	1
1.2. L'information textuelle.....	1
1.3. Problematique.....	2
1.4. Objectifs.....	4
1.5. Plan du memoire.....	4
<b>II. LES METHODES DE RECHERCHE D'UNE REPRESENTATION CONDENSEE.....</b>	<b>5</b>
II.1. Introduction.....	5
II.2. Presentation des methodes.....	6
II.2.1. Full text scanning.....	6
II.2.2. Inversion.....	7
II.2.3. Fichier signature.....	9
II.2.4. Clustering.....	11
II.2.5. Hashing multi-attributs.....	12
II.3. Conclusion.....	13
<b>III. METHODE DU FICHER SIGNATURE.....</b>	<b>14</b>
III.1. Methodes d'extraction de signature.....	14
III.1.1. Word signature.....	14
III.1.2. Superimposed coding.....	15
III.1.3. Compression.....	16
III.1.3.1. Bit-block compression.....	17
III.1.3.2. Run-length encoding.....	19
III.2. Choix de la methode.....	20

III.3. Choix de la longueur de la signature.....	21
IV. DOMAINE D'APPLICATION.....	28
IV.1. Introduction.....	28
IV.2. Domaine d'intervention de la methode du fichier signature.....	29
IV.2.1. Lors de la création.....	29
IV.2.2. Lors de la consultation.....	30
IV.2.3. Lors de la recherche.....	30
IV.2.3.1. Recherche par auteur.....	30
IV.2.3.2. Recherche par mot-clé.....	31
IV.2.3.3. Recherche par partie de titre.....	31
IV.3. Conception.....	32
IV.3.1. Introduction.....	32
IV.3.2. Analyse structuree.....	32
IV.3.2.1. Diagramme de flux de données.....	33
IV.3.2.2. Dictionnaire des données.....	35
IV.3.2.2.1. Flux des donnees (Data flow).....	36
IV.3.2.2.2. Rangement des données (Data store).....	37
IV.3.2.2.3. Les processus.....	37
IV.3.3. Modele structurée de solution.....	40
IV.3.3.1. Introduction.....	40
IV.3.3.2. Modele de specification de modules.....	41
IV.4. Implémentation.....	52
IV.4.1. Introduction.....	52
IV.4.2. Choix du SGBD.....	53

IV.4.3. Conclusion.....57

V. CONCLUSION.....60

BIBLIOGRAPHIE.....64