

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

INSTITUT NATIONAL D'INFORMATIQUE  
I.N.I

MEMOIRE DE FIN D'ETUDES

Pour l'obtention du diplôme d'Ingénieur d'Etat en Informatique

Option : SYSTEMES INFORMATIQUES

THEME

**SYSTEME DE GENERATION  
DE STRUCTURE HYPERTEXTE  
POUR LES BASES  
DE DONNEES BIBLIOGRAPHIQUES.**

PRESENTE PAR :

Melle KEDDARI Djalila  
Melle GHIDA Nassima

PROMOTEURS :

Mme A.EL-MAOUHAB  
Melle H.LABIOD

Centre d'accueil : CERIST  
(Centre de Recherche sur l'Information Scientifique et Technique)

- Promotion 1994 -

## RESUME

Le concept d'hypertexte s'adapte convenablement à l'obtention d'informations pertinentes dans le cas de bases de données bibliographiques volumineuses.

Parmi les étapes de constitution d'un système hypertexte, celle qui consiste à structurer les documents référentiels de manière à ce qu'ils soient exploitables par le système et qui constitue par conséquent une phase préparatoire importante.

L'objectif du projet est donc de développer un système qui **génère** à partir des données bibliographiques sous différents formats une structure adaptée au traitement hypertexte qui dispose de moyens de marquage.

Un besoin d'utilisation de normes est indispensable pour aboutir à une structuration logique des documents référentiels. La norme qui répond le mieux aux spécifications d'une telle structuration est la norme internationale de marquage de document SGML ( Standard Generalized Markup Language ). L'introduction du principe de marquage est faite dans le but d'adopter un modèle de structuration de données qui soit adapté au réseau hypertexte appelée HYPERSTRUCTUE et qui intègre la notion de noeuds et de lien entre les noeuds : concepts de base qui permettent de définir des hyperdocuments

Pour définir les liens et les noeuds dans le cas des bases de données bibliographiques, des techniques d'associations et de clusterisation sont utilisées.

## MOTS CLES

Base de données bibliographiques

Documents référentiels

Norme SGML

Indexation ,inversion ,clusterisation

Hypertexte ,hyperdocument, noeud , lien

# SOMMAIRE

INTRODUCTION.....	7
<b>Chapitre I :INTRODUCTION A LA DOCUMENTATION AUTOMATIQUE</b>	
I.1 INTRODUCTION .....	11
I.2 DEFINITIONS .....	11
I.2.1 Information .....	11
I.2.2 Document .....	12
I.3 BASES DOCUMENTAIRES .....	14
I.3.1 Définition d'une base documentake .....	14
I.3.2 Types de bases documentaires .....	14
1.4 SYSTEMES DOCUMENTAIRES .....	16
I.4.1 Definition .....	16
I.4.2 Langage documentake .....	17
I.4.3 Fonctions essentielles d'un système documentake.....	17
I.4.4 Critère de performance d'un système documentake .....	19
1.5 RECHERCHE DOCUMENTAIRE .....	20
I.5.1 Definition .....	20
I.5.2 Outils de la recherche dans une base documentake .....	22
CONCLUSION .....	23
<b>Chapitre II :ANALYSE DE DONNEES ,S.G.M.L , ET DOCUMENTATION</b>	-
II.1 ANALYSE DES MOTS ASSOCIES .....	25
II.1.1 Objectif de l'analyse des mots associés .....	25
II.1.2 Techniques d'analyse des mots associés .....	25
II.2 TAXINOMIE .....	27
II.2.1 Définition .....	27
II.2.2 Principaux concepts .....	27
II.2.3 Types de classification hiérarchique .....	28

II.3. S.G.M.L .....	<b>30</b>
11.3.1 Présentation de la norme SGML .....	30
11.3.2 Domaines d'application .....	31
II.3.3 Concepts de base de la norme SGML .....	32
11.3.4 Structure d'un document SGML .....	36
II.3.5 Outils nécessaires pour l'utilisation de SGML .....	37
 <b>Chapitre III : PRESENTATION DES SYSTEMES HYPERTEXTES</b>	
III.1 ORIGINE , EVOLUTION ET DEFINITION .....	<b>39</b>
III.1.1 Introduction .....	<b>39</b>
III.1.2 Historique .....	<b>40</b>
111.2 DOMAINES D'APPLICATION ET SYSTEMES EXISTANTS .....	44
111.2.1 Domaines d'application .....	44
III.2.2 Systèmes hypertextes existants .....	<b>45</b>
III.3. STRUCTURE D'UN SYSTEME HYPERTEXTE .....	38
III.3.1 Noeuds .....	<b>49</b>
III.3.2 Liens .....	<b>53</b>
111.4. EXPLORATION DU RESEAU .....	56
111.4.1 Les visionneurs .....	56
III.4.2 Recherche .....	56
111.5 ARCHITECTURE D'UN SYSTEME HYPERTEXTE .....	<b>58</b>
III.5.1 Modèle 1 : modèle de Campbell et Goodman .....	<b>58</b>
III.5.2 Modèle 2 : modèle de Dexter .....	<b>58</b>
III.6 AVANTAGES ET INCONVENIENTS .....	61
III.6.1 Avantage .....	61
III.6.2 Inconvénients .....	<b>61</b>
CONCLUSION .....	62

## **Chapitre IV : CONCEPTION DU SYSTEME DE GENERATION DE L'HYPERSTRUCTURE " HYPERGEN " A PARTIR D'UNE BASE DE DONNEES BIBLIOGRAPHIQUES**

IV.1 INTRODUCTION .....	<b>64</b>
IV.2 ARCHITECTURE DU SYSTEME DE GENERATION DE L'HYPERSTRUCTURE .....	<b>64</b>
IV.2.1 Modèle d'architecture .....	<b>64</b>
IV.2.2 Description de la chaîne de traitements .....	<b>65</b>
IV.3 FONCTIONNALITES DU SYSTEME DE GENERATION HYPERGEN.....	<b>67</b>
IV.3.1 Constructions des associations .....	<b>67</b>
IV.3.2 Clusterisation .....	<b>69</b>
IV.3.3 Balisage des clusters .....	<b>75</b>
IV.3.4 Hyperstructure .....	<b>78</b>
<b>Chapitre V : MISE EN OEUVRE DU SYSTEME DE GENERATION DE L'HYPERSTRUCTURE " HYPERGEN "</b>	
V.1 ARCHITECTURE MODULAIRE DU SYSTEME .....	<b>82</b>
V.2 DESCRIPTION DE LA CHAÎNE DE TRAITEMENT .....	<b>84</b>
V.2.1 Initialisation des données et construction des associations .....	<b>84</b>
V.2.2 Classification et construction des clusters .....	<b>86</b>
V.2.3 Construction des noeuds et des liens .....	<b>89</b>
V.2.4 Navigation dans l'hyperstructure .....	<b>92</b>
V.3 OUTILS DE VALIDATION DES STRUCTURES SGML .....	<b>104</b>
V.3.1 Lex et Yacc .....	<b>104</b>
<b>CONCLUSION</b> .....	<b>108</b>
<b>BIBLIOGRAPHIE</b> .....	<b>111</b>
<b>ANNEXES</b> .....	<b>115</b>
<b>Annexe A : NORME SGML</b> .....	<b>115</b>
<b>Annexe B : LEX ET YACC</b> .....	<b>127</b>