

**Ministère de l'Enseignement Supérieur et de la
Recherche Scientifique**
Centre de Recherche sur l'Information Scientifique et Technique
CERIST

MEMOIRE

pour l'obtention
du Diplôme de Post-Graduation Spécialisé (DPGS)
en Information Scientifique et Technique

***Contribution à la constitution d'un langage
d'indexation géographique :
Cas de l'ALGERIE***

par
Ahmed LIMAN

sous la direction du
Dr. Madjid DAHMANE

Alger, CERIST, 2000

DEDICACES

Je dédie le présent mémoire:

*à mes parents, et à mes beauxparents
à ma femme
à mes enfants, Djalel et Ikram*

et à tous ceux qui me sont chers

REMERCIEMENTS

Je voudrais exprimer ici ma reconnaissance a tous ceux qui, de près ou de loin, ont contribue' à l'élaboration de ce travail, je les assure de mon amitié en leur exprimant toute ma gratitude.

*A Monsieur **DAHMANE Madjid**, je voudrais qu'il trouve ici, le témoignage de ma reconnaissance pour avoir bien voulu accepter de diriger ce mémoire, et pour l'intérêt constant qu'il n'ajamais cessé de lui porter.*

*Mes remerciements vont également aux personnels de l'Agence National des Ressources en Hydraulique (ANRH), du Bureau National d'Etude et Développement Rural (BNEDER), du Centre de Recherche sur l'Information Scientifique et Technique (CERIST), de l'Institut National du Sol de l'Irrigation et du Drainage (INSID), de l'Office National de Ge'ologie (ONG), de l'Institut National de la Vulgarisation Agricole (INVA), sans pour autant oublier de nommer Monsieur **TALEB Said**, Directeur Général de l'I.N.V.A, Monsieur **KEFIL Rabah**, Directeur des Réseaux d'Information et de la Documentation Agricole a l'I.N.V.A, Mademoiselle **KHOURI Radia**, ingénieur agronome au service traitement, pour l'aide apportée a ce travail par leur efforts généreux manifestés ici et la, pour le rendre effectif.*

A tous, unefois de plus Merci.

Ahmed LIMAN

INTRODUCTION

L'Algérie connaît un taux de croissance démographique parmi les plus élevés du monde , De plus, avec ses 3 milliards de \$US/an d'importation en denrées alimentaire, qui représente 25% des importations totales du pays , le secteur agricole devrait prendre une part prépondérante dans l'activité économique .

Mais cette agriculture est loin d'être compétitive , les raisons sont multiples :

- Incertitude liées au climat : Pluviométrie capricieuse
- Démographie galopante
- Pauvreté des sols
- Effets pervers des pesticides
- Utilisation des techniques culturales non appropriés, etc.

Un autre facteur déterminant dans la non performance de l'agriculture est la carence et la non circulation de l'information pertinente et efficace, permettant de minimiser les effets des phénomènes adverses .

L'information agricole est recherchée par un public varié allant des producteurs agricoles aux chercheurs et décideurs politiques, en passant par les agents de vulgarisation ainsi que les spécialistes de la planification agricole et rurale.

En effet, selon les précisions de la F.A.O les pays en développement devraient pouvoir augmenter leur productivité agricole d'une moyenne de 3.8% à 4.1% à partir de l'an 2000 grâce à l'utilisation des connaissances scientifiques et techniques disponibles.

En Algérie, l'information agricole est notamment la documentation écrite, et est produite et traitée dans des centres de documentation et des bibliothèques d'institutions techniques dépendant du ministère en charge de l'agriculture et des services connexes tels que forêts, pêche, environnement, hydraulique, etc.

Ces différentes structures travaillent avec des outils documentaires de la F.A.O ou de structures étrangères.

Avec la généralisation de l'utilisation de l'outil informatique, on assiste à une banalisation de l'ordinateur au sein de ces structures de traitement documentaires. A côté de l'outil informatique on retrouve désormais d'autres outils documentaires conçus par des institutions de coopérations internationales telle que la F.A.O, et rarement d'outils documentaires d'initiative locale .

C'est dire aujourd'hui, que toute cette panoplie de nouvelles technologies, sont désormais en place au sein des unités de documentation agricole qui pour la plupart n'ont plus grand chose à envier à des unités de documentation similaires dans les pays développés.

Devant ces faits, il nous a paru nécessaire de mettre en place un langage d'indexation qui doit refléter les subdivisions géographiques nationales pertinentes, et destinés à couvrir la terminologie d'un domaine de connaissance spécifique, afin de s'en servir pour une meilleure interprétation des connaissances scientifiques et techniques dans le domaine de l'agriculture et des sciences connexes .

Les perspectives d'avenir d'un secteur économique tel que l'agriculture dépendra largement de ce dispositif permanent et pertinent de recueil d'information , et d'interprétation des données .

La qualité des services que nous proposerons demain dépend très largement de ce vocabulaire structuré, normalisé, et dynamique qu'est le langage d'indexation et de sa capacité d'interprétation des documents en langage naturel dans un système de langage contrôlé plus cohérent.

PROBLEMATIQUE

L'importance d'un langage d'indexation dépend largement de sa capacité d'interprétation de données, destinée à mettre en place un dispositif permanent et pertinent de recueil d'information.

Pour surmonter les difficultés créées, lors de la recherche de l'information, par l'ambiguïté des mots du langage naturel, l'une des solutions consiste à :

- Normaliser la terminologie utilisée pour la description géographique.

Cette étude a pour but de définir les descripteurs géographiques donc les identificateurs relatifs aux subdivisions géographiques pertinentes, basés sur des critères administratifs, agroclimatiques et physiographiques.

L'indexeur doit avoir toujours présent à l'esprit que la finalité de l'indexation est la récupération, dans une base de données, des informations répondant avec pertinence aux demandes des utilisateurs.

Ce langage d'indexation s'adresse aux indexeurs travaillant pour les systèmes documentaires nationaux et étrangers tels que dans notre cas les systèmes AGRAL et AGRIS, c'est à dire aux personnes responsables de la description des subdivisions géographiques d'un document.

Ce langage d'indexation est un outil fondamental pour l'amélioration des systèmes documentaires déjà existant, et va permettre principalement de :

- développer des bases de données bibliographiques locales (95 à 98% des cas)
 - transférer des données vers les sources externes (input AGRIS/CARIS)

Ce langage d'indexation va être développé autour du thesaurus AGROVOC, et du plan de catégorisation AGRIS, et va permettre en particulier d'enregistrer des termes d'indexation autres que ceux provenant du thesaurus AGROVOC, et va compléter cette indexation afin de mieux décrire les réalités locales pour les utilisateurs locaux et externes.

Dans ce langage, le descripteur ou l'identificateur géographique désigne la description du document comme terme d'indexation.

La mise en pratique de ce langage d'indexation dans les systèmes documentaires AGRAL et AGRIS devrait améliorer le travail des indexeurs et parfaire la recherche en information documentaire.

Les concepts relatifs à l'environnement et à la localisation sont spécialement importants pour les bases de données agricoles. La localisation est retenue dans l'analyse de toutes les unités traitant d'essai en champ, de situation économique, sociale, légale, etc...

Il faut garder à l'esprit qu'une localisation spatiale peut également signifier des conditions physiques particulières qui devront être spécifiées dans l'analyse.

L'indexeur doit analyser le contenu des unités, puis exprimer le sujet dans le(s) langage(s) autorisé(s) par le système pour lequel il prépare les données.

La qualité de l'indexation dépend de deux séries de facteurs, les uns liés aux langages d'indexation et les autres à l'indexeur lui-même.

Bien évidemment, la qualité des langages documentaires mis à la disposition de l'indexeur contribue à la qualité de l'indexation.

Ce langage d'indexation a un rôle dans l'élaboration et l'utilisation des bases de données. Il est l'un des facteurs majeurs de la qualité et de la fiabilité des informations fournies par les bases de données, c'est à dire de leur valeur pour les utilisateurs.

C'est dans cet objectif que notre mémoire va porter sur la contribution à l'élaboration d'un langage d'indexation géographique local.

Résumé

Les concepts relatifs à la localisation sont spécialement importants pour les bases de données agricoles.

Cette localisation définie par des identificateurs géographiques, donc par un langage d'indexation géographique, est l'un des facteurs majeurs de la qualité et de la fiabilité des informations fournies par les bases de données.

Cette outil d'indexation normalise développer autour de thésaurus et de classification va compléter l'indexation afin de mieux décrire la bibliographie nationale.

Les perspectives d'avenir d'un secteur économique tel que l'agriculture dépendra largement de ce dispositif permanent et pertinent de recueil d'information, et d'interprétation de données.

Mots clés

LANGAGE D'INDEXATION GEOGRAPHIQUE / IDENTIFICATEUR GEOGRAPHIQUE / SUBDIVISION ADMINISTRATIVE / SUBDIVISION CLIMATIQUE/ SUBDIVISION PHYSIOGRAPHIQUE/ THESAURUS/

Notice bibliographique

Contribution à l'élaboration d'un langage d'indexation géographique : cas de l'Algérie. / Liman, A. Institut national de la Vulgarisation Agricole (Algérie) These (Diplôme de post-graduation spécialisée en Information Scientifique et Technique).

Jun 2000. 91p..Numerous tables ; numerous graphs ; 31 ref.. Annexe.

SOMMAIRE

Introduction

Problématique

1- Présentation de l'Institut National de la Vulgarisation Agricole (I.N.V.A)	01
11/ <i>Dénomination</i>	01
12/ <i>Dare de création</i>	01
13/ <i>Statut</i>	01
14/ <i>Missions et prérogatives statutaires</i>	01
15/ <i>Objectifs techniques</i>	02
16/ <i>Potentiel humain</i>	04
17/ <i>Organisation interne</i>	04
18/ <i>Organigramme</i>	06
2- Présentation de la Direction des Réseaux d'Information et de la Documentation Agricole (D.R.I.D.A.)	07
21/ <i>Missions et prérogatives statutaires</i>	07
22/ <i>Organisation interne</i>	09
221/ Service bibliothèque	09
2211/ Activités	09
2212/ Fond documentaire	09
2213/ Personnel	09
2214/ Matériel	10
222/ Service Traitement documentaire	10
2221/ Activités	10
22211 Outils et langage documentaire	10
22212 Personnel	10
22213 Matériel	11
223/ Service Réseau documentaire	11
2231/ Activités	11
2232/ Logiciels	11
2233/ Personnel	11
2234/ Matériel	11
23/ <i>Organigramme et Activités</i>	13

24/ Circuit des informations et leurs exploitations _____	14
3- Présentation des systbmes documentaires Agris et AGRAL _____	15
31/ <i>Présentation du système documentaires AGRIS</i> _____	15
311/ Situation actuelle _____	15
312/ Domaine d'activité _____	16
313/ Rôle de la F.A.O _____	17
314/ Role des pays membres _____	17
32/ <i>Présentation du système documentaire AGRAL</i> _____	18
321/ Situation actuelle _____	18
322/ Domaine d'activité _____	18
323/ Rôle _____	19
3231/ Animation du reseau _____	19
32321 Contrôle, traitement et maintenance _____	19
324/ Rôle des membres du reseau _____	20
33/ <i>Fonctionnement du système AGRIS et AGRAL</i> _____	20
331/ Preparation du bordereau d'entrée des données AGRIS _____	21
3311/ Manuel d'opération pour la description bibliographique (guide de catalogage) _____	22
A/ <i>Description bibliographique</i> _____	24
B/ <i>Description chronologique des zones</i> _____	24
C/ <i>Matériel auxiliaire</i> _____	24
D/ <i>Prksentation des éléments de données bibliographiques</i> _____	24
3312/ Guide d'indexation _____	32
A/ <i>Introduction</i> _____	32
B/ <i>Indexation et recherche d'information</i> _____	32
C/ <i>Indexation locale</i> _____	33
D/ <i>Les langages d'indexution</i> _____	33
D1/ Historique _____	33
D2/ Les classifications _____	34
D3/ Les langages analytiques _____	35
E/ <i>Normalisation terminologique du langage d'indexation</i> _____	36
F/ <i>Langage d'indexution et langage naturel</i> _____	36
G/ <i>Les différents types de langages d'indexation</i> _____	31
G1/ Les classifications _____	31
G11/ Définition _____	31
G12/ Classifications encyclopédiques _____	38
G13/ Classifications spécialisées _____	38
G14/ Propriétés _____	39

G15/ Notation des classifications	39
G16/ Critères de division	40
G17/ Niveau de spécificité	40
G18/ Codification des classes	40
G19/ Cas d'une classification Plan de catégorisation AGRIS	40
G191/ Introduction	40
G192/ Domaine d'AGRIS et d'AGRAL	41
G193/ Structure du système de catégorisation AGRIS	41
G194/ Spécificité des catégories AGRIS	42
G195/ Les codes des catégories	42
G196/ Les différents éléments du plan	42
G197/ Plan de catégorisation AGRAL/AGRIS	44
G198/ Agrindex et Agral-index	48
G199/ Recherche bibliographique	48
G2/ Les thesaurus	49
G21/ Historique	49
G22/ Définition	49
G23/ Présentation du thesaurus	50
G24/ Les différents types de mots	50
G25/ Les descripteurs	51
G26/ Les non-descripteurs et la relation d'équivalence	51
G27/ La structure du thesaurus et la relations entre Descripteurs	52
G271/ Les relations hiérarchiques	52
G272/ Les relations associatives	52
G28/ Spécificité d'un thesaurus	52
G29/ Cas d'un thesaurus: AGROVOC	52
G291/ Présentation	52
G292/ Domaine thématique d'AGROVOC	53
G293/ Multilinguisme	53
G294/ Structuration	54
a) Ordre alphabétique des termes et jeu de caractère	54
G295/ Descripteurs, non-descripteurs, notes d'application et termes permutés	55
a) Descripteurs	55
b) non-descripteurs	56
c) Notes d'application	58
d) Les termes permutés	59
G296/ Forme des descripteurs et des non-descripteurs	59
G297/ Precordination	60
G298/ Les relations de termes	62
a) Les relations hiérarchiques (TG: terme générique, TS: terme spécifique)	62
b) Les relations associatives (TA: terme associé)	64
c) Les relations de préférence (EP: employé pour, EM: employé)	64
G299/ Traitement particulier de certaines catégories de terme	65
a) Les noms taxonomiques des plantes	65

b) Les noms taxonomiques des animaux _____	66
G300/ Mise à jour d'AGROVOC _____	67
332/ Logiciel CDS/ISIS _____	68
3321/ <i>Introduction</i> _____	68
3322/ <i>Aperçu du logiciel CDS/ISIS</i> _____	69
3323/ <i>Fonction du logiciel CDS/ISIS</i> _____	70
3324/ <i>Structure de la base de données</i> _____	71
4- Constitution du langage d'indexation _____	72
<i>i Introduction</i> _____	72
41/ <i>Choix du langage d'indexation</i> _____	72
42/ <i>Règles d'établissement du langage d'indexation géographique local</i> _____	73
/Avant propos _____	73
421/ <i>Objet et domaine d'application</i> _____	73
43/ <i>Rôle du langage d'indexation dans la fonction documentaire</i> _____	74
44/ <i>Choix des termes descripteurs ou identificateurs</i> _____	75
441/ <i>Identificateur</i> _____	75
442/ <i>Critère de choix des identificateurs</i> _____	75
443/ <i>Choix de la forme des identificateurs</i> _____	76
444/ <i>Choix de la représentation des identificateurs</i> _____	76
4441/ <i>Abréviation et sigles</i> _____	16
44421 <i>Orthographe</i> _____	16
4443/ <i>Termes compost avec des mots de liaison</i> _____	76
44441 <i>Jeu de caractères</i> _____	77
4445/ <i>Ponctuation et caractère spéciaux</i> _____	77
445/ <i>Note explicative ou note d'application</i> _____	77
45/ <i>Les relations entre descripteurs</i> _____	78
451/ <i>Relations d'équivalence</i> _____	78
452/ <i>Relation hiérarchique</i> _____	79
4521/ <i>Relation générique</i> _____	19
45211 <i>Relation partitive</i> _____	80
453/ <i>Relations associatives</i> _____	80
46/ <i>Construction du langage d'indexation</i> _____	81
/Préliminaire _____	81

461/ Collecte des termes significatifs du langage naturel, candidat descripteurs _____	81
4611/ Méthode analytique _____	81
4612/ Méthode synthétique ou globale _____	81
4612/ Combinaison des deux méthodes _____	82
462/ Collecte des termes sur les subdivisions administratives _____	82
463/ Collecte des termes sur les subdivisions climatiques _____	82
464i Collecte des termes sur les subdivisions physiographiques _____	86
465/ Choix et évaluation des descripteurs _____	86
4661 Etablissement de relations _____	86
4661/ Regroupement thématique _____	86
466 <i>li</i> Regroupement par facettes _____	87
47/ Présentation du langage d'indexation _____	
4711 Introduction du langage d'indexation _____	87
472/ Types de présentation _____	88
4721/ Présentation alphabétique _____	88
4722/ Présentation alphabétique permutée _____	88
47231 Présentation thématique _____	88
4724/ Présentation hiérarchique _____	89
4725/ Présentation graphique _____	89
48/ Essai et mise à jour du langage d'indexation _____	89
481/ Essai _____	89
4821 Mise à jour _____	90

Conclusion