

E.N.S.S.I.B.
Ecole nationale supérieure
des sciences de l'information
et des bibliothèques

DPSSIB
Diplôme professionnel supérieur en sciences de
l'information et des bibliothèques

Rapport de recherche bibliographique

Vocabulaires Contrôlés et OPACs :
les problèmes de l'intégration dans un système
utilisant des vocabulaires contrôlés multiples

Nadia BOUCHAIB
Sous la direction de :
Madjid IHADADJENE
Doctorant au Centre d'Etude et
et de Recherche en Sciences de
l'Information (CERSI)



1996-1997

E.N.S.S.I.B.
Ecole nationale supérieure
des sciences de l'information
et des bibliothèques

DPSSIB
Diplôme professionnel supérieur en sciences de
l'information et des bibliothèques

Rapport de recherche bibliographique

Vocabulaires Contrôlés et OPACs :
les problèmes de l'intégration dans un système
utilisant des vocabulaires contrôlés multiples

Nadia BOUCHAIB
Sous la direction de :
Madjid IHADADJENE
Doctorant au Centre d'Etude et
et de Recherche en Sciences de
l'Information (CERSI)

1997
DPS
BIB
02

1996-1997



Résumé :

Par ce travail, nous avons voulu exposer les problèmes de l'intégration de plusieurs vocabulaires contrôlés au sein de systèmes de recherche documentaire, tels que les catalogues en ligne : OPACs. Nous nous sommes intéressés aux différentes approches de résolution de systèmes utilisant au niveau de l'accès au sujet, plusieurs vocabulaires contrôlés. Dans ce sens, les interfaces de recherche ont été améliorées pour une meilleure pertinence au niveau de l'accès au sujet. Nous présentons en exemple, le système NOTIS version 5. 0, une expérimentation satisfaisante. Enfin, nous avons également tenter d'identifier et localiser quelques uns de ces systèmes.

Descripteurs :

Vocabulaire contrôlé / accès sujet / recherche documentaire en ligne / thesauri multiples / catalogue automatisé / OPACs / autorité matière / vedette matière / thesaurus / classification / liste autorité / compatibilité / système intégré / interface utilisateur / CDD / CDU / LCSH / Mesh / NOTIS / DOBIS / UMLS

Abstract :

With this present work, we tried to examine many problems consisting in multiples controlled vocabularies integration for subject searching system, such as Online Catalogs : OPACs. We present different systems resolution approaches using more than one controlled vocabularies for a subject access search. In this way, users searches interfaces had been ameliorated. We present also, NOTIS 5. 0 system exemple such as one succesful experimentation. For ending, we tried to identify and localize some of these systems.

Descriptors :

Controlled vocabulary / subject access / Online Public Access Catalog / OPACs / Online information retrieval / Multiple thesauri / multithesauri / multi- bases system / user interface / subject heading / subject authority / classification / thesaurus / authority file / DDC / UDC / LCSH / MeSH / compatibility / convertibility / integrated system / vocabulary switching system / VSS / intermediate lexicon / DOBIS / NOTIS / Unified Medical Language System / UMLS

Table des matières

1 APPROCHE ET MÉTHODOLOGIE DE RECHERCHE.....	3
1.1 STRATÉGIE DE RECHERCHE	3
1.1.1 Définition de mots clés et descripteurs	3
1.2 EXPLORATION ET UTILISATION D'OUTILS DE RECHERCHE.....	4
1.3 EQUATIONS DE RECHERCHE UTILISÉES :	5
1.3.1 Recherche dans les outils imprimés	5
1.3.2 Interrogation en ligne via le serveur DIALOG	7
1.3.3 Interrogation de bases Cédéroms (CD-ROM) : INSPEC et LISA.....	8
1.3.4 Interrogation via Internet :	10
1.4 ESTIMATION ET COÛT.....	11
2 ANALYSE ET SYNTHÈSE DU SUJET	11
2.1 PROBLÉMATIQUE	12
2.2 LE SUPPORT INFORMATIQUE POUR LES VOCABULAIRES CONTRÔLÉS MULTIPLES.....	14
2.3 LES PROBLÈMES DE L'INTÉGRATION.....	15
3. LES APPROCHES POUR ASSURER LA COMPATIBILITÉ DANS DES SYSTÈMES À VOCABULAIRES MULTIPLES	16
3.1 Lancaster et Smith :	16
3.2 Mandel et le système de la Bibliothèque du Congrès :	17
4 LES MOYENS DE RÉOLUTION DU PROBLÈME DE L'INCOMPATIBILITÉ :	18
4.1 L'intégration et la fusion des vocabulaires :	19
4.2. Les systèmes de langages intermédiaires :	20
4.3 Macrothesaurus et microthesaurus :	20
4.4 Classification clustering :	20
4.5 Les Systèmes de recherche de catalogues intégrés :	22
4.6 L'Index sujet unifié UMLS : Unified Medical Language System.....	24
4.7 NOTIS (Northwestern Online Total Integrated System) :	24
4.8 IDENTIFICATION DES SYSTÈMES DE RECHERCHE :	29
5. CONCLUSION.....	31
3. BIBLIOGRAPHIE.....	32
3.1 GÉNÉRALITÉS	32
3.2 INTÉGRATION ET FUSION DE VOCABULAIRES.....	34
3.2.1 Classification clustering	43
3.2.2 UMLS.....	45

1 Approche et méthodologie de recherche

Le sujet pose une problématique qui est l'accès au sujet au sein des systèmes de catalogues en ligne OPACs utilisant au niveau de la recherche et de l'indexation plusieurs vocabulaires contrôlés.

Cette problématique bien précise, que nous développons dans la deuxième partie de notre travail, nécessite la méthodologie de recherche suivante :

D'abord de situer le sujet dans son contexte : le sujet nous renvoi au problème plus précis qui est l'incompatibilité des langages ou vocabulaires contrôlés au niveau de la recherche au sujet. Notre première approche du sujet nous a amené à le situer dans le domaine des sciences de l'information ou sciences documentaires sous les thèmes qui sont : la recherche de l'information, l'accès au sujet, les catalogues automatisés, les vocabulaires contrôlés, les systèmes documentaires, la conception d'interface de recherche dans les OPACS utilisant des vocabulaires de recherche multiples notamment, les vocabulaires contrôlés qui sont les thesaurus, les classifications, les listes d'autorités ou listes de vedettes matières...etc. Une deuxième approche nous a permis l'identification et la localisation de ces systèmes. ce sont là les aspects essentiels de notre sujet.

1.1 Stratégie de recherche

La stratégie qui s'impose est de rechercher les différentes approches sur le sujet à partir des descripteurs définis : les descripteurs matière et les noms d'auteurs et spécialistes de la question. Une deuxième étape est d'identifier et localiser les différents systèmes fonctionnant sur l'intégration et la fusion des vocabulaires contrôlés multiples, donc interroger par les noms des systèmes.

1.1.1 Définition des descripteurs

Pour les différentes étapes de la recherche, nous avons défini des descripteurs d'accès et descripteurs d'interrogation. en raison de la nature technique du sujet, qui nous a amené au début de notre recherche, à constater l'abondance de documents publiés dans cette langue, traitants de notre sujet, ces derniers ont été définis en langue anglaise. Donc, ce choix s'est imposé de lui même, puisque nous avons utilisé un pourcentage important de documents et d'outils en anglais.

Descripteurs spécifiques liés au sujet :

Controlled vocabulary / Subject access / Online Public Access Catalog / OPACs / Online information retrieval / Multiple thesauri / multithesauri / Multi- bases system / User interface / Subject headings / subject authority / Classification / Thesaurus / Authority file / DDC / UDC / LCSH / MeSH / compatibility / convertibility / integrated system / vocabulary switching system / VSS / intermediate lexicon / Unified Medical Language System / UMLS

Vocabulaire contrôlé / Accès sujet / Recherche documentaire en ligne / Thesauri multiples / Système de bases de données multiples / Interface utilisateur / catalogue automatisé / OPACs / Vedette matière / Vedette autorité / Classification / Thesaurus / liste autorité / compatibilité / système intégré / CDD / CDU / LCSH / MeSH.

Descripteurs liés au noms d'auteurs et spécialistes du sujet :

Aitchison, R. / Chamis, A. / Chan, L. M. / Cimino, J. J. / Lancaster, F. W. / Larson R. R. / Mandel, C. A. / Niehoff, R. T. / O'Brien A. / Olson, T. / O'Kane, K. C. / Sanjeevi, A. / Strawn, G. L.

Descripteurs liés au noms des systèmes de recherche dans les OPACs :

DOBIS / ETHICS / NOTIS / ORION / TERM / UMLS

1.2 Exploration et utilisation d'outils de recherche

Compte tenu de la spécificité et la technicité du sujet, nous avons exploré et utilisé les outils de recherche de trois types. Les outils papier : monographies, périodiques et publications en série. Les outils électroniques : catalogues en ligne et bases bibliographiques en ligne et sous forme de cédéroms (CD-ROM), ainsi qu'une interrogation via le réseau Internet.

Dans ce sens, nous avons utilisé les outils support papier existants au Centre de documentation et à la bibliothèque de l'ENSSIB, à savoir, les dossiers de presse, les monographies, des publications en série et périodiques spécialisés anglo-saxonnes en majorité. Nous citons deux publications en série qui nous ont permis d'avancer dans notre approche : Annual Review in Information Science and Technology de l'ASIS (American Society of Information Science) et Encyclopedia of Library and Information Science. Nous avons dépouillé un certain nombre de périodiques spécialisés : Cataloging and Classification Quarterly, Journal of Documentation, Journal of Librarianship and Information Science, Library Trends, Libri (International Library Review), Online review. Nous avons également utilisé les périodiques spécialisés localisés à la bibliothèque municipale de Part Dieu : Information Technology and Libraries, International Classification, Program... etc

La thèse de Sanjeevi, que nous avons demandé par le prêt entre bibliothèque, nous a été d'une grande utilité par sa richesse d'informations. Quant aux monographies nous avons entamé notre approche du sujet à partir des ouvrages de Chamis, Chan, Lancaster, de Mandel et d'autres auteurs spécialistes du sujet.

Quant la recherche dans les outils électroniques, nous avons utilisé l'OPAC de la bibliothèque, les bases de données spécialisées, répondants à notre sujet sous forme de cédéroms : LISA (Library and Information Science Abstracts), complétée par l'année en court sous forme papier. Pour sa couverture du domaine de l'informatique et sciences de l'ingénieur, nous avons consulté INSPEC (Information Services in Physics, Electrotechnology, Computers and Control : Institution of Electrical Engineers : IEE Royaume Uni) à la bibliothèque de l'INSA. L'aspect informatique de notre sujet nous a amené au choix de cette dernière source.

L'interrogation en ligne dans les bases de données spécialisées : LISA, ISAbstracts, INSPEC et PASCAL, via le serveur DIALOG, a été effectué lors d'un exercice en séance de travaux dirigés (RDI) à l'Urfist. Pour les deux premières bases bibliographiques citées, le choix est évident puisque LISA et ISA de couvrent dans le domaine des sciences de l'information et de la bibliothéconomie au niveau universel. En raison du temps limité de cette interrogation en ligne, nous n'avons pu affiner notre recherche lors de cette interrogation.

Enfin nous avons effectué une interrogation via le réseau Internet qui a été des plus fructueuse parce qu'elle nous a permis d'identifier quelques systèmes de recherche en ligne que nous citerons plus loin et d'interroger le catalogue de la US National Library of Medicine NLM et de Ameritech library services. Pour cette dernière interrogation, le site utilisé est : Altavista par le moteur de recherche : « advanced search » ^{ou} à partir des adresses électroniques suivantes :

- <http://www.Web-search.Com/home21.html>;
- <http://home.netscape.com/internet-search.html>;
- <http://www.als.ameritech.com/>;
- <http://www.index.nlm.nih.gov/cgi/htsearch> ;
- <http://www.cs.uni.edu/~okane> ou okane@uni.edu;
- <http://www-camis.S...tsheet.umls95.html>

1.3 Equations de recherche utilisées :

Nous avons procédé à une stratégie de recherche de type booléen à plusieurs niveau : c'est à dire que nous avons utilisé autant de combinaisons possible que de descripteurs définis. Il est certain que, dans certain système de recherche nécessitant l'utilisation de descripteurs composés, nous avons employé ces descripteurs sans fragmentation.

1.3.1 Recherche dans les outils imprimés

Encyclopedie of Library and Information Science : nous avons cherché à partir de l'index sujet (vol. 47) par les points d'accès suivants

Points d'Accès	Résultats
Access to sources of knowledge	1
Access points Computerized catalog	1
Design of controlled vocabularies	1
Indexing languages	1

On- line subject searching	1
Vocabularies	1
On- line information systems	2
Controlled vocabulary	1
On- line indexing	1
Vocabulary structure	1
Subject headings	1
Subject indexes	1
Vocabularies for online subject searching	1

Annual Review in Information Science and Technology : nous avons cherché à partir de l'index sujet (vol. 26, 1991) aux points d'accès suivants :

Accès	Résultats
Online public access catalogs (OPAC)	1
Classification	1
Subject access	1
Subject searching	1
Vocabulary control	1
Online thesauri	1
Subject analysis	1
Thesauri voir aussi	1
Thesaurus	2
Information retrieval	7
Online searching	3
Thesauri	1
for online catalogs, computer- readable dictionaries	1
Use with classification schemes	1

Online keyword access to public information voir [OKAPI]	
NOTIS / LUIS	1

1.3.2 Interrogation en ligne via le serveur DIALOG

4 : INSPEC 1983- 1997/ janv

61 : LISA (LIBRARY & INFOSCI) 1969- 1996 / oct.

144 : Pascal 1973- 1996 / dec.

202 : Information Science Abs. 1966- 1996 / dec.

? S Controlled (w)vocabular ??

Set	File	Items	Description
	4	166	
	61	250	
	144	201	
	202	234	

S1 851 controlled(w)vocabular ??

? S OPAC ?? or (Online(w)Public(w)Access(w)Catalog ?)

? S1 and S2

4 : INSPEC

166	S1
596	S2
6	S1 AND S2

144 : Pascal

201	S1
3493	S2
10	S1 AND S2

61 : LISA

250	S1
1224	S2
10	S1 and S2

202 : Information Sciencee Abs.

234	S1
513	S2
9	S1 AND S2

? S integration and system ??

144 : Pascal

41719 INTEGRATION

61 : LISA

1865 INTEGRATION

1711129 SYSTEM ??
 11058 INTEGRATION
 AND SYSTEM ??

42265 SYSTEM ??
 873 INTEGRATION
 AND SYSTEM

202 : Information Science Abs
 4188 INTEGRATION
 2963394 SYSTEM ??
 3193 INTEGRATION
 AND SYSTEM ??

4 : INSPEC
 64760 INTEGRATION
 108736 SYSTEM ??
 33354 INTEGRATION
 AND SYSTEM ??

S4 48478 INTEGRATION AND SYSTEM ??

? S3 and S4

INSPEC

6 S3
 33354 S4
 0 S3 AND S4

LISA

10 S3
 873 S4
 0 S3 AND S4

PASCAL

10 S3
 11058 S4
 0 S3 AND S4

ISA

9 S3
 3193 S4
 0 S3 AND S4
 S5 0 S3 AND S4

TOTAL S1 and S2 = 35

? ts3/5/1- 35, dont 11 résultats pertinents

1.3.3 Interrogation de bases Cédéroms (CD-ROM) : INSPEC et LISA

INSPEC	Résultats	Pertinence
# 1 Multiple	7925	
# 2 thesaur?	73	
# 3 # 1 and # 2	4	4
# 4 Controlled vocabular?	18	10

# 5	vocabulary control	2	0
# 6	Compatibilit?	1134	5
# 7	6 and # 2	0	
# 8	online public access catalog?	9	
# 9	OPAC	28	
# 10	8 or 9	30	
# 11	subject heading?	13	3
# 12	# 11 and 10	0	0
# 13	subject access	1	1
# 14	classification	22	7
# 15	# 13 and 14	1	1

LISA	Résultats	Pertinence
1. SU = compatibility	36	
2. SU = thesauri	864	
3. CS = 1 and 2	3	3
4. SU = controlled vocabulary	15	5
5. SU = controlled vocabularies	1	1
6. CS = 1 and 4	0	0
7. SU = controlled	40	
8. SU = vocabulary	28	
9. CS = 7 and 8	15	5
10. CS = 1 and 9	0	0
11. Ab = vocabularies	136	
12. CS = 7 AND 8	54	
13. SU = OPACs	9	
13. Ab = online catalog	12	
14. Ab = integrated information system	798	2
15. CS = 10 and 13	3	3
16. Ab = switching	284	
17. CS = 8 and 16	10	5
18. Ab = cluster based retrieval	20	6
19. Ab = metathesaurus	14	2

20. SU = metatheasurus	2	2
21. Ab = authority files	190	6
22. Ab = classification	4748	
23. CS = 21 and 22	11	4
24. SU = authority control	27	3
25. SU = multiaccess	1	1
26. Ab = umls	22	7
27. SU = ORION	11	5
28 AU = Congreve	5	1
29. AU = Hildreth	22	1
30. AU = Lancaster	182	4
31. AU = Markey	41	4
32. AU = Niehoff	7	4

1.3.4 Interrogation via Internet :

à partir du moteur de recherche « Altavista »

Equations	Résultats	Pertinents
Compatibility and convertibility and controlled near vocabularies	1	1
integrated near vocabularies	40	4
Multibase near integration	9	
Multithesaurus or multiple near thesauri	76	
Multiple near vocabularies and online near catalog	7	7
Multiple near vocabularies and OPAC	4	4
Unified near index and OPAC	68	4
Chamis near Alice	5	3
O'Kane near Kevin near Charles	34	9

Olson near tony	2	2
Sanjeevi near Amba	2	2
Strawn near Gary near L.	7	2

1.4 Estimation et coût

Le coût et le temps de l'interrogation effectué le 09 janvier, via le serveur Dialog, à l'URFIST est estimé à :

Sub account : BOUCHAIB

\$2.97 0.198 Hrs File 4

\$0.006 6 types in format 5

\$0.00 6 types

\$2.97 Estimated cost File 4

\$0.48 0.032 Hrs File 61

\$0.00 10 types in Format 5

\$0.00 10 types

\$0.48 Estimated cost File 61

\$3.45 0.230 Hrs File 144

\$0.00 10 types in Format 5

\$0.00 10 types

\$3.45 Estimated cost File 144

\$0.59 0.039 Hrs File 202

\$0.00 9 types in format 5

\$0.00 9 types

\$0.59 Estimated cost File 202

One search, 4 files, 0.500 Hrs File OS

\$3.00 NETGNS

\$10.49 Estimated cost this search

\$11.19 Estimated total session cost 0.533 Hrs

2 Analyse et synthèse du sujet

Préambule

Dans le monde de l'accès en ligne, nous assistons de plus en plus à un développement considérable des systèmes de recherche d'information. Pourtant la multiplicité des vocabulaires contrôlés utilisés dans les systèmes de recherche crée des complications affectant réellement deux types de services d'information. D'une part les services bibliographiques, qui offrent des centaines de bases de données diverses à travers une interface d'interrogation unique. D'autre part sont affectés les systèmes de bibliothèques intégrées ou les systèmes de catalogues de bibliothèques en ligne : **Online Public Access Catalogs**.

Ces systèmes de recherche en ligne qui produisent des index automatiquement, offrent comme moyen d'accès au sujet le langage naturel en utilisant les mots figurants dans le titre, le résumé, le texte libre et les langages conventionnels obéissants à un contrôle de syntaxe des descripteurs retenus, appelés les vocabulaires contrôlés ou langages d'indexation.

L'environnement de ces sources d'information, c'est à dire les catalogues de bibliothèques en ligne offre donc une multiplicité de vocabulaires. Néanmoins cette multiplicité de vocabulaires contrôlés n'est pas sans difficulté au niveau d'abord, du contrôle de ce vocabulaire ensuite au niveau de l'accès ou la recherche au sujet.

Des approches, des techniques, des expérimentations mais aussi des systèmes fonctionnels de recherche d'information ont été conçus comme alternatives de résolution aux complications provoqués par cette incompatibilité au niveau des langages.

Pour aborder le sujet, nous partirons de la problématique suivante :

2.1 Problématique

La recherche documentaire en ligne dans les catalogues informatisés de bibliothèques ou dans les bases de données multiples implique l'utilisation de vocabulaires contrôlés multiples ou de plusieurs thesauri. Mais l'utilisation de vocabulaires et de thesauri multiples implique à son tour l'identification des termes dans ces différents thesauri. Les termes constituant ces vocabulaires contrôlés devraient être semblables les uns aux autres et substituables les uns aux autres.

Sans nul doute, le problème majeure qui reste sans solution définitive dans les systèmes automatisés de recherche en ligne, est de pouvoir développer les mécanismes nécessaires pour résoudre les problèmes d'incompatibilités. Cette dernière, engendrée par **l'intégration de vocabulaires multiples** coexistants dans un même système. l'objectif donc, est d'améliorer l'accès au sujet dans les systèmes de catalogues en ligne par la conception d'interfaces de recherche documentaire améliorées. Il est évident que ce problème d'incompatibilité entre plusieurs vocabulaires, comme nous l'avons évoqué, gêne le transfert des données enregistrées

entre systèmes utilisant des langages contrôlés différents et la recherche dans des bases de données utilisant des thesaurus différents.

Des approches de résolution du problème de l'intégration ont été exposés puis expérimentés. L'amélioration de ces systèmes et de leurs interfaces de recherche suscitent un intérêt particulier chez les gestionnaires de bibliothèques, les concepteurs et spécialistes d'OPACs. Mais la solution au problème n'est pas encore définitive lorsque nous avons des systèmes de recherche en ligne qui comportent des dizaines de bases de données utilisant une multiplicité de vocabulaires contrôlés, le cas de DIALOG pour les systèmes bases de données et NOTIS pour les systèmes de recherche de catalogues en ligne, par exemple.

Nous nous intéressons dans cette recherche bibliographique aux problèmes de l'intégration de plusieurs vocabulaires contrôlés dans un même système de recherche documentaire en ligne pour améliorer l'accès au sujet. En d'autres termes, au problème de l'incompatibilité de ces différents langages lorsqu'ils coexistent dans un même système, intervenant sérieusement au niveau de la pertinence des résultats d'une recherche à partir de la reconnaissance du sujet.

Il s'agit de savoir, quelles sont les sources de cette incompatibilité engendrant des difficultés au niveau de l'intégration de plusieurs vocabulaires dans le même système de recherche en ligne et quels sont les moyens de cette intégration.

Egalement, quels sont les problèmes rencontrés lors de l'interrogation ou de la recherche.

Enfin, Il s'agit également d'identifier les systèmes de recherche qui supportent plusieurs vocabulaires ou plusieurs thesaurus.

Nous citons ci-après un exemple simple de système intégrant deux vocabulaires :

Si dans un système nous avons : A thesaurus, et

B thesaurus

Soit dans ce système A et B sont deux thesaurus établis selon les même modèles et dont les descripteurs peuvent s'additionner et s'employer pour indexer et rechercher les documents de concepts équivalents.

Alors, A et B ne peuvent être compatibles que si dans un même domaine, et dans le même système, les documents indexés avec le thesaurus A peuvent être retrouvés avec le thesaurus B et Si les documents indexés avec le thesaurus B peuvent être retrouvés avec le thesaurus A. Cette compatibilité ou relation de réciprocité implique qu'au niveau des désignations et des définitions de leurs descripteurs il existe une correspondance.

2.2 Le support informatique pour les vocabulaires contrôlés multiples

Posons nous maintenant la question : Qu'est ce que les vocabulaires contrôlés ? C'est un ensemble limité de termes qui doivent être utilisés pour représenter le concept matière des documents. Ce type de vocabulaires peut être : une liste d'autorité matières ou liste de vedettes matières, une classification, un thesaurus ou simplement une liste de mots clés et phrases approuvées. Ces même vocabulaires sont utilisés pour la recherche au sujet.

Pour se faire, un index sujet est généré par la « machine- readable » dans laquelle est stocké le catalogue automatisé et dont les informations sont interprétées directement par l'ordinateur. Cet index qui utilise un langage spécifique « Index Language » est une partie du système de recherche d'information. Il peut être sous forme de liste alphabétique ou classificatoire des termes retenus lors de l'indexation : vedettes matières, ou encore une variété des deux : vedettes matières, indices de classification. Une interface de recherche à partir de laquelle un menu d'aide à l'utilisateur propose des renvois d'orientation utilisant ces même vocabulaires pour l'accès au sujet.

Depuis les années 1960, des vocabulaires contrôlés, notamment, les thesaurus, avaient été développés séparément pour les catalogues spécialisés, les index ou les services d'information, lorsque les différences entre leurs vocabulaires n'apparaissaient pas pour poser un problème à leurs utilisateurs au niveau de l'interrogation (ZENG, 155). Mais l'évolution importante des systèmes de recherche de l'information provoque un changement.

Comme ces catalogues ont été automatisés, le défi était de munir le support informatique du plus de sujets possible. Ce support informatique de vocabulaires contrôlés impliquait un support pour trois fonctions essentielles qui sont : la construction du thesaurus, le contrôle d'autorité matière, la recherche au sujet.

Aux Etats Unis notamment, différentes approches, mais aussi des projets ne cessent d'être élaborés ces dernières décennies. L'objectif, comme nous l'avons évoqué auparavant, est de rendre ces systèmes de recherche en ligne compatibles à la multiplicité de vocabulaires qui ne cessent de poser problème et de les munir d'interfaces de recherche conviviale et accessible à tout utilisateur. D'abord au niveau de l'indexation même : la tendance étant l'exhaustivité et l'universalité d'un domaine général de connaissances comportant des domaines plus spécifiques. Aussi au niveau de la recherche au sujet : la pertinence et la précision de l'accès au sujet à partir d'une interface de recherche aisée et d'un système de recherche d'information performant.

Cette incompatibilité se traduit bien sûr, automatiquement au niveau du transfert des données enregistrées entre ces différents systèmes de vocabulaires qui sont : les thesaurus les classifications universelles, les listes d'autorités matières et les langages naturels. L'alternative était de munir les systèmes de recherche d'outils lexicaux appelés : les thesaurus de recherche que nous énumérons comme suit :

- ◆ Les thesaurus fusionnés et intégrés : ces thesaurus offrent une terminologie plus riche et une plus grande variété de corrélations entre les termes que les vocabulaires indépendants. Ce point est développé avec plus de détails dans la partie (3) de notre analyse.

- ◆ Les réserves de termes : ce sont des fragments de stratégies de recherche, conservés en vue d'un usage ultérieur. Les « réserves » ou « facettes » de recherche sont en général composées d'un certain nombre de termes précédés de l'opérateur booléen « ou » dans le listing. Pour être retrouvées, ces stratégies de recherche nécessitent au niveau de l'indexation des renvois réciproques.
- ◆ Les thesauri traditionnels élargis : pour une recherche dans une base de données en langage naturel et traitant du même domaine, on peut utiliser les thesauri traditionnels. C'est à dire, utiliser plusieurs thesaurus au lieu d'un seul. Il est possible d'avoir recours à la fois aux descripteurs et aux non- descripteurs de ces thesaurus pour la recherche. Les thesaurus traditionnels comportant des index permutés et des représentation hiérarchiques et systématiques aident à trouver plus rapidement les termes pertinents. Il est possible d'élargir les thesaurus traditionnels pour effectuer une recherche en langage naturel par l'augmentation du nombre de termes renvoyant au descripteur le plus proche.
- ◆ Les liste de termes de bases de données : c'est des listes générées par la machine dans un ordre alphabétique, contenant les mots significatifs de la base de données, avec indication de fréquences, mais sans signalement des relations.
- ◆ Les listes de synonymes : ce sont des listes de termes et leurs équivalents qui peuvent être substitués à un terme au cours d'une recherche en ligne.

2.3 Les problèmes de l'intégration

Posons-nous désormais la question : quels sont les problèmes qui peuvent intervenir lors de l'intégration de plusieurs vocabulaires contrôlés ? D'après Chamis, ces complications existent dans l'utilisation de thesaurus pour un enregistrement sujet, au niveau conceptuel, terminologique et classificatoires. (CHAMIS, 44).

- Les problèmes conceptuels incluent l'ambiguïté de compréhension et la complémentarité du vocabulaire;
- Les problèmes terminologiques incluent le degré de l'équivalence des termes, le degré de pré-coordination et post-coordination des termes;
- Les problèmes classificatoires incluent la spécificité et l'exhaustivité des classes de termes, les termes reliés ou les relations d'apparentées

De ce qui précède, nous pouvons dire que l'incompatibilité se situe donc, à deux niveaux : au niveau de la spécificité des termes indexés ; et au niveau de l'exhaustivité du domaine indexé, par exemple :

- ◆ Un système X peut comporter une terminologie précise et détaillée, alors qu'un système Y, principalement des termes généraux;
- ◆ Un thesaurus X peut omettre certains aspects du domaine, tandis qu'un thesaurus Y, peut en couvrir tous les aspects ;

- ◆ Un système X peut utiliser des termes pré-coordonnés ou composés là où un autre Z exprimera les mêmes concepts en combinant des termes indépendants;
- ◆ Le choix des formes préférentielles parmi les synonymes et les quasi-synonymes diffère d'un thesaurus à un autre;
- ◆ Les hiérarchies des thesaurus peuvent avoir des structures différentes et être plus ou moins développées : c'est à dire qu'au niveau des relations entre les termes, dans un thesaurus X, des niveaux hiérarchiques peuvent apparaître, alors qu'elles peuvent ne pas figurer dans un thesaurus Y.
- ◆ D'un thesaurus X à un thesaurus Y, les termes en relation associative peuvent varier, puisque le choix d'un terme associé peut être influencé par le domaine d'intérêt ou le point de vue propre à l'organisme pour lequel est élaboré le thesaurus. Une variété de méthodes avait été proposée pour résoudre ces problèmes et faciliter la compatibilité ou l'intégration entre les langages d'indexation.

3. Les approches pour assurer la compatibilité dans des systèmes à vocabulaires multiples

Nous ne pouvons exposer les différentes approches et expérimentation sans donner un aperçu de l'évolution qui a donné naissance à toutes ces approches dans l'environnement automatisé de systèmes de recherche documentaire que nous connaissons actuellement.

Dans les années 1970, l'absence de compatibilité du vocabulaire d'indexation posait un problème réel au niveau de l'interrogation qui apportait aux utilisateurs des résultats insatisfaisants pouvant augmenter le coût de la recherche.

L'Unesco, publie une étude de faisabilité sur le système des sciences de l'information universel UNISIST (UNISIST, 33). La principale recommandation à l'intention des scientifiques est " d'initier les associations de sciences de l'information au besoin de développer de meilleurs outils pour le contrôle et la conversion des langages d'indexation et des langages naturels dans les sciences et technologie".

D'autres conférences avaient inclus dans leur programme le problème de l'intégration et de l'incompatibilité des langages documentaires dans les systèmes de recherche à vocabulaires multiples. Aussi la possibilité de normalisation et de conception de systèmes de langages unifiés (UNESCO, 30, 31, 32)

Plusieurs conférences internationales ont depuis concentré leurs travaux sur les intérêts de la compatibilité. A présent, vingt cinq ans après le rapport UNISIST, les systèmes de recherche en ligne se sont beaucoup améliorés et des expérimentations ont abouti.

3.1 Lancaster et Smith :

Dans un rapport du Programme générale d'information de l'Unesco, Lancaster, F.W. (LANCASTER, 87) et Smith, L.C (SMITH, 90), avaient exposé leurs approches permettant d'assurer la compatibilité du vocabulaire en quatre catégories :

- ◆ Mapping
- ◆ Le lexique intermédiaire : switching language ou vocabulaire fusionné;
- ◆ Le vocabulaire intégré : integrated vocabulary
- ◆ Les microthesauri et les macrothesauri ou macrovocabulaires.

3.2 Mandel et le système de la Bibliothèque du Congrès :

Dans un rapport préparé pour les services de traitement de la Bibliothèque du Congrès, Carol A. Mandel (MANDEL, 96), exposait en 1987 la multiplicité de thesaurus dans les systèmes bibliographiques des bibliothèques en ligne nord américaines. Ces bibliothèques avaient normalisées leurs vocabulaires contrôlés avec l'utilisation des vedettes matières (subject headings) de la Bibliothèque du Congrès : Library of Congress subject headings (LCSH), séparé les catalogues utilisant les thesaurus spécialisés habitués à certains sujets. Le système supporte des vocabulaires de différents sujets spécialisés. Les quatre approches de base données par Mandel, à l'époque , pour fournir l'accès aux bases de données indexées avec des vocabulaires différents sont :

a/ Les enregistrements séparés : Séparer chaque enregistrement bibliographique représentant un vocabulaire. les bases de données utilisant différents thesaurus sont cherchées séparément;

b/ Les vocabulaires fusionnés ou mélangés : il s'agit de produire un « **lexique intermédiaire** », Vocabulary Switching System VSS : les vocabulaires sont mélangés ou fusionnés dans un seul enregistrement bibliographique. les termes de la totalité des vocabulaires sont recherchés ensemble par la recherche au sujet;

c/ Les vocabulaires intégrés " Integrate vocabulary " : les techniques utilisées pour ^{associer} relié les différents thesaurus peuvent être appliquées à leurs tour pour développer les structures " syndetic "; c'est à dire pour fournir une structure d'accès intégrée permettant d'aider la recherche dans un environnement de bases de données multiples;

d/ La navigation de vocabulaires multiples à travers les techniques assistées par ordinateurs. Les caractéristiques des termes sont développées dans les interfaces intelligentes de recherche en ligne et peuvent être désignées pour l'aide à partir de bases de données multiples.(MANDEL 91).

e/ Les Microthesauri et Macrothesauri ou macrovocabularies.

Mandel décrit le contrôle d'autorité effectué dans la liste d'autorités matières ou vedettes matières LCSH : Library of Congress subject Headings. Cette liste superstructure menue de codes de compatibilité lesquels sont utilisés dans quatre microthesauri qui ont été développés comme satellites autour de LCSH. Ces microthesauri sont :

- ◆ Le vocabulaire d'indexation législatif : the legislative indexing vocabulary (LIV)
- ◆ Le thesaurus graphique de la bibliothèque du congrès : the LC thesaurus of graphic materials (TGM)
- ◆ Les vedettes matières de la littérature enfantine : subject headings for children's literature
- ◆ Le thesaurus d'art et architecture : Art and Architecture thesaurus (AAT)

Ces quatre domaines sont étroitement liés et peuvent être intégrés dans le système lorsque le thesaurus est maintenu comme une liste d'autorité et indexé puis relié aux enregistrements bibliographiques de recherche.

Dans ce sens chaque terme LIV, par exemple, est assigné au code de compatibilité LCSH comme suit :

LC - le terme est le même dans LCSH

LCX - le terme est " voir" dans LCSH

LCC - le terme est similaire à un terme dans LCSH

LCD - les caractères sont identiques mais le sens est différent.

Weintraub (WEINTRAUB, 149), décrit l'utilisation de LCSH et les vedettes matières de MeSH : Medline subject Headings dans un même catalogue. Les vedettes matières sont indexées séparément et chaque index est accessible à partir d'un menu principal. Inévitablement les utilisateurs choisissent, parfois une option de recherche inappropriée, mais d'après Weintraub, l'ensemble des réactions sont favorables .

4 Les moyens de résolution du problème de l'incompatibilité :

De ce qui précède, nous nous posons la question de savoir quels sont les moyens ou techniques pour résoudre le problème de l'incompatibilité, c'est dire rendre ces systèmes de recherche plus compatibles et supportent la multiplicité des vocabulaires ? Une variété de méthodes avait été proposée pour résoudre ces problèmes et faciliter l'intégration entre les vocabulaires d'indexation et de recherche :

- ◆ La Matrice de compatibilité conceptuelle de Dahlberg;
- ◆ Le Microthesaurus et Macrothesaurus de Thompson;
- ◆ Le thesaurus intégré de Aitchison;
- ◆ Les modèles de données de Soergel;
- ◆ Le thesaurus intégré et le système de vocabulaire fusionnée de Niehoff.

Trois moyens pour résoudre les problèmes d'incompatibilité sont exposés respectivement par Niehoff et Aitchison :

4.1 L'intégration et la fusion des vocabulaires :

L'expérimentation la plus connue est le système de fusion de vocabulaire, un mécanisme de fusion de sujet automatiquement à titre expérimental pour chercher dans des bases de données multiples par un langage naturel unique développé par Niehoff et d'autres à la Battelle Columbus Laboratories (USA).

Cette alternative tend à résoudre le problème de l'incompatibilité en fournissant les correspondances entre des thesaurus couvrant le même domaine. la technique consiste à fusionner les termes en les classant par ordre alphabétique dans une liste unifié. Cette fusion se déroule en deux phases :

La première phase consiste à comparer les descripteurs et les entrées par ordre alphabétique. Quant à la deuxième, elle consiste à intégrer dans une seule entrées les termes en correspondance et leurs relations sémantiques pour dégager les cohérences.

La création du vocabulaire intégré a encore besoin de beaucoup de travail intelligent. L'approche améliorée est d'utiliser les systèmes automatisés pour le contrôle d'autorité et pour fusionner le vocabulaire de recherche (NIEHOFF, 113, 114).

Deux exemples de vocabulaires fusionnés :

4.1.1 BRS/TERM :

La base de données BRS/TERM est un exemple de vocabulaire fusionné de termes contrôlés dans le domaine des sciences du comportement et des sciences sociales interrogeable en ligne. Ce système donne aussi les équivalents en langage naturel des termes contrôlés.

4.1.2 VSS : Vocabulary switching system :

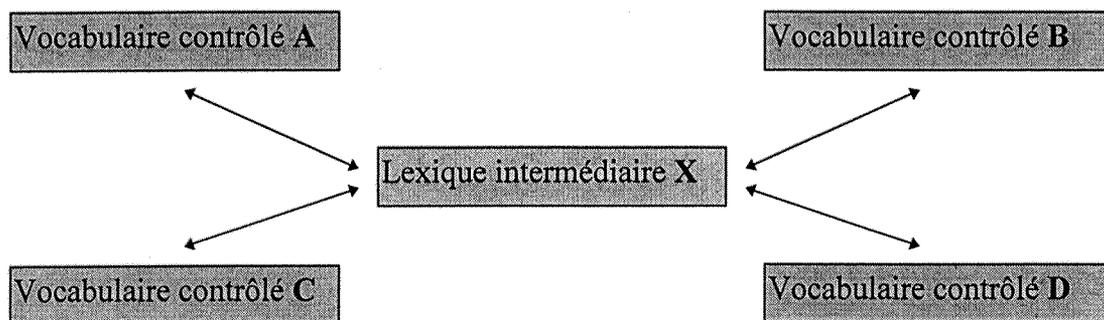
Est une base de données expérimentale développée aux Laboratoires Battelle Colombus en Ohio, décrite par Niehoff (NIEHOFF, 113, 114). Le champ couvert par le VSS concernait à l'origine l'énergie et des thèmes apparentés, mais celui-ci s'est étendu à quatre domaines : les sciences physique, les sciences de la vie, les sciences sociales et le commerce. Le nombre de vocabulaire fusionnés est passé de six à quinze. En réponse à l'entrée d'un terme, le système signale les correspondances exactes et des synonymes des termes génériques, spécifiques et associés aux correspondances des mots : c'est à dire de tous les termes composés et de toutes les locutions dans lesquelles le mot apparaît. C'est le système de langage intermédiaire.

4.2. Les systèmes de langages intermédiaires :

Comme nous l'avons exposé précédemment, cette technique consiste en la création d'un "lexique intermédiaire" permettant d'établir des passages entre le système et les langages d'indexation, aussi bien les classifications que les thésaurus.

Si dans un système, A, B, C et D sont des langages d'indexation ou vocabulaires. X est le lexique intermédiaire, donc :

- le lexique intermédiaire X assure le passage entre
- le langage d'indexation A et les langages B, C et D entre
- le langage B et les langages A, C, et D entre
- le langage C et les langages A, B et D et ainsi de suite



Il a été conclu que pour obtenir un maximum d'efficacité dans la recherche au sujet, le lexique intermédiaire doit avoir la même spécificité que le langage le plus développé. L'étude a révélé que la conversion du langage B dans le langage D au cours d'une recherche dans la base de données D était aussi performante que le langage D à lui seul. (HORSNELL, 79).

Le 4.3 Macrothesaurus et microthesauri :

un microthesaurus est un thesaurus spécialisé qui s'adapte à la structure hiérarchique d'un thesaurus plus général : le macrothesaurus, qui comporte l'ensemble de microthesaurus. Il peut être encyclopédique, indépendamment des microthesaurus qui sont compatibles avec lui. Il peut exister aussi sous une forme amalgamée associant la structure du macrothesaurus encyclopédique et les microthesaurus spécialisés. Nous avons cité un exemple de macrothesaurus et de microthesauri au début de cette analyse (l'approche Mandel), c'est le système de la Bibliothèque du Congrès. LCSH, est une liste d'autorité superstructurée, c'est le macrothesaurus, LIV, SHCL, TGM, et AAT sont des microthesauri.

4. 4 Classification clustering :

C'est une technique de recherche basé sur le "groupage de termes en grappes selon la classification". D'après Le dictionnaire des sciences de l'information et technology « Dictionary of Information Science and Technology », le « **Clustering** » est le groupage des items dans une base de données dont les éléments du grappe présente des ressemblances les uns les autres et des non ressemblances des autres grappes.

Dans les systèmes de recherche, les questions « items » dans un groupe sont souvent recherchées ensemble (en même temps) dans la réponse vers la question. Pour chaque grappe la question composée peut être générée est utilisée comme la base de la recherche du grappe.

Il existe deux classes générales des méthodes du clustering :

1. Les méthodes hiérarchiques qui produisent des groupes « nichés » et ;
2. Les méthodes non hiérarchiques (tels que les méthodes de liaison unique), qui produisent une couche unique de grappes qui peuvent être dédordés : plusieurs questions « items » peuvent être produit dans plus qu'un grappe. Ce groupage est utilisé en même temps pour la liaison des questions et pour la classification automatique des documents.

4.4.1 O'Kane et Sanjeevi

L'utilisation des techniques de clustering ou « groupage de termes en forme de grappe » pour la recherche de documents pertinents dans les catalogues en ligne avait été rapporté par **Sanjeevi Amba**, dans une étude (Thèse de PHD) (SANJEEVI, 187). Celui-ci examine la faisabilité du « Vector - Space Model » et expose l'approche du « Cluster- Based » : le groupage de termes en grappes qui était utilisé pour associé les termes à partir de deux thesauri. Une méthode unique d'association de termes « Single Link Method » est décrite en exposant le problème comme suit :

Pour chaque terme du **thesaurus (1)** est associé un groupe de termes dans le **thesaurus (2)** comme si que la recherche basé sur un groupe de termes du **thesaurus (2)** est aussi efficace que la recherche utilisant le terme original du **thesaurus (1)**.

Un groupe de documents indexés à partir de termes contenus dans deux thesaurus : le **thesaurus ERIC** et le **Psychological Index Terms** étaient identifiés et ces documents ayant des relations commune formaient la base de cette étude. Les co- occurrences

d'**ERIC** et les descripteurs de **PSYC** dans le groupe de ces documents qui ont des relations communes étaient utilisé pour former des groupages de termes en grappes :

« Clusters ». Ces groupages en grappes « Clusters » contiennent les termes à la fois d'**ERIC** et les termes de **PSYC** pour assurer la liaison en même temps que se forment

les groupes. L'efficacité de liaison est contrôlée par la validation de chaque groupe

« cluster » sur la base de critères préétablis utilisant la précision et le rappel, deux mesures standards dans la recherche de l'information.

voir
p.20
corrige

Il a été prouvé, d'après cette expérimentation, que ces procédures développées peuvent être utilisées pour lier ou associer plusieurs thesaurus dans un même système.

Sanjeevi énumère les avantages de cette approche comme suit :

- ◆ La relation effectuée est fondée sur les termes du thesaurus désignés comme descripteurs ;
- ◆ La flexibilité ou la souplesse dans la formation des groupes « clusters » est possible ;
- ◆ Les changements de terminologie ou dans les politiques d'indexation n'affecteront pas les procédures de base déjà développées ;
- ◆ Enfin les problèmes de résolution de l'incompatibilité entre les thesaurus sont évités.

Néanmoins, il a été remarqué que :

- ◆ Les termes qui sont dans le « Cluster » ou qui apparaissent en même temps ne font pas toujours preuve de relations thesaurales ;
- ◆ Un terme d'un thesaurus n'est pas toujours associé à un seul terme du deuxième thesaurus ;
- ◆ Les termes équivalents du point de vue lexical, ne se produisent pas toujours en même temps.

4.5 Les Systèmes de recherche de catalogues intégrés :

Aux Etats Unis et au Canada, le contrôle d'autorité est rigoureusement développé dans les systèmes de catalogues de bibliothèques intégrées notamment, dans les systèmes catalogues de bibliothèques universitaires tels que **ORION, WLN, NLC/DOBIS, UTLAS, Geac BPS, NOTIS et Carlyle.**

Ces systèmes peuvent fournir le contrôle d'autorité pour plus d'un vocabulaire sujet séparé logiquement. Généralement ils ont deux fonctions majeurs. La première est d'identifier la liste source de sujet, alors que l'autre est de valider les vedettes matières d'après les vocabulaires sélectionnés. La tendance actuelle est d'établir des liens entre les classification systématiques et les listes de vedettes matières.

Des expériences de recherche en ligne sur les OPACs : **SULIRS, NELVYL, NOTIS, SCORPIO** dans différentes bibliothèques universitaires des Etats Unis et de la Bibliothèque du congrès ont conduit à l'analyse des capacités du LCSH (Library of Congress Subject Headings) en format lisible par machine pour servir d'outil principal d'aide à la recherche. Quatre points sont exposés :

- ◆ Adéquation entre le vocabulaire des questions et le vocabulaire contrôlé ;
- ◆ Connaissance de ce dernier par les utilisateurs ;
- ◆ Proposition en ligne de termes associés, renvois d'orientation des termes du LCSH vers les vedettes matières retenues dans les enregistrements ;

Proposition de tactiques de recherche sous forme d'organigrammes.

D'après les résultats d'une enquête par questionnaire produite au Congrès de l'ALA (American Library Association) en 1987, il a été conclu que la plupart des systèmes de recherche offrent la manipulation de termes d'autorités tel que LCSH et MeSH. L'interface utilisateur de ces systèmes comportent en outre comme aides aux utilisateurs lors d'une recherche dans les OPACs : l'affichage des termes autorisés, les renvois d'orientation, une terminologie diverse pour voir et voir aussi, les notices d'autorités et des messages à l'écran. (NYE, 116)

En Europe, des systèmes comme DOBIS (Dortmunder Bibliotheks System) de la bibliothèque universitaire de Zurich et ETHICS, un système de recherche développé par la bibliothèque suisse de l'Institut Fédéral de Technologie de Zurich (ETH Bibliothek), sont des exemples de systèmes de recherche dans des catalogues intégrés.

A titre d'exemple nous présentons le système ETHICS tel qu'il a été décrit par HUG et WALSER (HUG, 81). Ce système comprend un OPAC, le contrôle d'autorité, la description et le catalogue sujet. Neuf bibliothèques techniques suisse participent au réseau ETHICS. Son interface de recherche est en allemand et en français et la base de données couvre environ 55.0000 entrées. Ces deux systèmes ont conçu leurs vedettes matères de manière à ce q'elles concordent avec les indices de la classification en usage.

La recherche dans ETHICS peut être au mots du titre (recherche au texte libre), par le mot sujet, le vocabulaire contrôlé et le feuilletage ou la navigation dans la liste alphabétique des auteurs et titres. Pour les recherches au sujet, les entrées au catalogue sont classifiés avec les codes de la CDU. L'accès au sujet est basé sur un registre de termes associés en allemand, anglais et en français avec leurs codes CDU.

L'utilisateur n'a pas besoin de connaître les codes de la CDU. Il peut effectuer sa recherche au sujet à partir de termes. La recherche au sujet et ses entées apparentées, la recherche au concepts reliés, aux termes génériques et spécifiques et la combinaison de plusieurs concepts sont démontré.

En Europe toujours, la Commission XII- B3 de la Communauté Economique Européenne a tracé les grandes lignes de procédure pour la création d'un Metathesaurus que la commission a appelé : « Base de données de thesaurus multifonctions et multilingues ». Une liste de près de mille thesaurus utilisé dans le monde a déjà été compilé.

En Chine, le projet : Chinese Classified- Thesaurus, initié par la Bibliothèque Nationale de Chine (la bibliothèque de Pékin), vise à élaborer un vocabulaire intégré basé sur la classification standard nationale : Chinese Library Classification

(CLC) et la thesaurus standard national recommandé : Chinese Thesaurus (CT). Le nouveau vocabulaire fournit pour chacune des classes de CLC des descripteurs correspondants de CT et vice versa (ZENG, 155).

4.6 L'Index sujet unifié **UMLS** : Unified Medical Language System

Initié par la National Library of Medicine (NLM) en 1986, ce projet consiste à développer un système de langage médical unifié. L'intention est d'aider les professionnels et chercheurs dans le domaine des sciences médicales et de la santé, à chercher et pouvoir intégrer les informations biomédicales à partir d'une variété de sources. Ce système intègre quatre vocabulaires : les vedettes matières de sciences médicales MeSH, la nomenclature systématique de médecine SNOMED, le Current Medical Information and Terminology CMIT et le système de recherche de l'institut national du cancer PDQ. Un métathésaurus est utilisé pour emmagasiner les concepts et les termes médicaux dans une forme canonique de chacun des vocabulaires multiples existants. Chaque enregistrement Meta - 1 contient trois types d'information : les faits de base, les relations et les données d'usage.

4.7 **NOTIS** (Northwestern Online Total Integrated System) :

Comme nous l'avons déjà exposé, les problèmes au niveau de l'accès au sujet surviennent lorsque deux ou plus de systèmes de recherche au sujet sont combinés dans le même index sujet.

La nécessité de normalisation de plusieurs vocabulaires contrôlés coexistants dans un même catalogue ou même système de recherche s'est concrétisée par l'élaboration d'index de sujet unifié.

Dans le domaine des sciences médicales, nous avons identifié le système **NOTIS**. Ce système combine la liste des vedettes matières de la Bibliothèque du Congrès **LCSH** et la liste de vedettes matières des sciences médicales, **MeSH** : Medical Subject Headings.

Nous avons jugé utile de présenter le système **NOTIS** pour son évolution et son efficacité au niveau de la recherche selon ses auteurs. Ainsi nous donnons une description détaillée avec des exemples d'affichage de ce système de recherche.

Olson, Tony (OLSON, 215), cite les problèmes les plus importants engendrés par la cohabitation de MeSH et LCSH, dans sa présentation de **NOTIS version 5. 0** qui a la capacité de recherche sur un sujet universel dans le catalogue en ligne. D'après Olson, ces problèmes sont :

1. La Duplication des entrées index pour le même enregistrement parce que les vedettes matières sont identiques dans les deux systèmes, exemple :

Abdominal pain (mal abdominal) est vedette matière à la fois dans MeSH et dans LCSH;

2. Les contradictions entre les renvois dans un système et les vedettes matières dans les autres systèmes, exemple :

Cancer est vedette matière retenue pour le système dans LCSH, mais c'est une référence de renvoi dans MeSH. Alors que **Neoplasms** est vedette matière pour MeSH, et devient renvoi dans LCSH;

3. Des résultats de recherche sont incomplètes et qui prêtent à confusion parce que des concepts ou descripteurs identiques sont représentés différemment dans les deux systèmes, exemple :

La vedette matière : **History of medicine, 17 th century** et

la vedette matière : **Medicine - History - 17 th**, représentent les même descripteurs;

4. Dans deux systèmes, pour une même orthographe de la vedette matière (descripteur) nous avons deux sens différents, exemple :

Crowns (couronne) renvoi à :

Dental appliance (appareil dentaire), dans MeSH. Alors que dans LCSH :

Crowns, renvoi à **Royal regalia** (emblème royal)

Depuis 1990, des travaux ont été élaborés pour trouver des solutions à ce problème de vocabulaire multiple par la Northwestern University et la Division américaine de programme technologique des bibliothèques de collèges.

Pour les systèmes d'OPACs dans le domaine de l'éducation, les solutions ont été développées pour les problèmes (1) et (3) que nous avons cité précédemment.

Quant au problème (4), les solutions sont du ressort des services de conception et de mise à jour des systèmes d'accès au sujet.

Deux principaux objectifs sont à atteindre :

1. L'utilisation des logiciels existants ou développer de nouveaux logiciels pour résoudre les problèmes (1) et (2) cités auparavant;

2. Affichage des systèmes MeSH et LCSH à partir des premières correspondances entre les vedettes matières de MeSH et LCSH, ensuite faire rentrer toutes les correspondances dans la liaison de l'entrée (spécialement le champ 750 et 788) champs dans les enregistrements d'autorité. Les résultats peuvent être utilisé pour résoudre le problème (3).

La solution du problème (1) est incluse dans le logiciel de NOTIS version 5. 0. Quant une vedette matière identique à MeSH et LCSH est présente dans la notice bibliographique (fig. 1), les programmes qui génèrent les indexes d'OPAC supprime simplement une des entrées de l'affichage. Afin de résoudre le problème (2), un nouveau programme a été élaboré à Northwestern University : ce programme était additionné au logiciel NOTIS 5. 0. Ce caractère est appelé " **Dynamic Conflict Resolution** ". Dans le contexte de l'index sujet unifié de NOTIS , lorsque le renvoi est incompatible à la vedette établie, la Dynamique résolution d'incompatibilité change de référence de niveau de recherche : " **Search Under** " dans une référence : **Search Also Under** ". Par exemple :

Clinical Pharmacology est vedette établie dans LCSH, mais un renvoi dans MeSH. La recherche de sujet dans l'**index unifié** de " Pharmacologie clinique " devrait produire un affichage d'index qui devrait inclure " Search under reference ", " Search also under reference " et trois notices bibliographiques (fig 3).

L'utilisateur pourrait rencontrer des confusions. Mais le Dynamic Conflict Résolution (fig 4) est là pour dire que l'information additionnée peut être trouvée sous une autre vedette.

Par ailleurs, il a été noté que c'est un bon exemple pour faire augmenter le résultat de la recherche à partir de la recherche au sujet universel.

Par la réorientation ou le détournement de la recherche vers la vedette matière MeSH, **Pharmacology clinical**, 73 notices bibliographiques sont récupérées.

Figure 1

MARC record containing identical (after normalization)

MeSH AND LCSH headings for ABDOMINAL PAIN

Search request : S=ABDOMINAL PAIN

NUC cat

Search results : 12 entries found

Subject index

Figure 2

ABDOMINAL PAIN

- 1 * For information on this heading, type 1

ABDOMINAL PAIN

* Search also under

- 2 ACUTE ABDOMEN
- 3 COLIC
- 4 DIGESTIVE SYSTEM DISEASES
- 5 GASTROINTESTINAL DISEASES
- 6 PAIN

ABDOMINAL PAIN

- 7 ABDOMINAL PAIN (1979) (HS)
- 8 BASIC AND CLINICAL ASPECTS OF CHRONIC ABDOMI (1993)
- 9 CHRONIC PELVIC PAIN IN WOMEN (1981) (HS)
- 10 GUT REACTIONS (1980) (HS)
- 11 PLAIN X RAY IN THE DIAGNOSIS OF THE ACUTE ABDOMEN

(1971)

Continued on next screen

Figure 3

Universal subject search for CLINICAL PHARMACOLOGY

Without dynamic conflict resolution

Search request : S=CLINICAL PHARMACOLOGY

NU cat

Search result : 5 entries found

Subject index

Figure 4

CLINICAL PHARMACOLOGY

1 * Search also under the equivalent heading : PHARMACOLOGY, CLINICAL
CLINICAL PHARMACOLOGY

2 * Search also under : DRUGS—TESTING

CLINICAL PHARMACOLOGY

3 MELMON AND MORRELLIS CLINICAL PHARMACOLOGY
(1992)

4 PHARMACOLOGIC BASIS OF ANESTHESIOLOGY (1994)
CLINICAL PHARMACOLOGY—UNITED STATES

5 FORGING A WOMENS HEALTH RESEARCH AGENDA (1991)

La plupart du temps, les efforts consacrés à ce problème étaient utilisés pour atteindre le deuxième objectif, l'affichage des systèmes MeSH et LCSH.

Le reste de cette présentation résumera les résultats du projet d'affichage et commente comment les résultats peuvent être utilisés. L'affichage était réalisé en deux phases. Dans une première phase, les correspondances entre les vedettes MeSH et LCSH avaient trouvé

l'utilisation d'une combinaison du traitement de la machine et le traitement humain. La méthodologie employé dans cette partie du projet a été présenté dans d'autres rencontres.

Plusieurs exemples des correspondances sont démontrées dans (fig 5- 6). Dans la majorité des cas les correspondances s'effectue une à une, mais il existe des exceptions qui sont démontrées en (fig 6). Dans certains cas (exemple a) deux ou plus de vedettes dans un système étaient affiché à une vedette unique dans un autre système .

Dans d'autres cas (exemple b) la vedette dans un système était équivalente à deux vedettes matières, dans un autre système elles sont combinées en utilisant l'opérateur booléen " and ".

La vedette LCSH : **Bone densitometry** est équivalente dans les vedettes MeSH à :

Bone and Bones and Densitometry utilisé ensemble dans le même enregistrement.

Olson note, qu'il y avait peu de cas dans lequel le concept était représenté par la vedette principale dans un système et sous vedette " subheading " dans un autre système.

Figure 5 et 6

- a/ MeSH : Nuclear Reactors
 LCSH : Nuclear reactors
- b/ MeSH : Pharmacology, Clinical
 LCSH : Clinical pharmacology
- c/ MeSH : Art Therapy—in infancy & childhood
 LCSH : Art therapy for children
-

Les correspondances entre les vedettes MeSH et LCSH sont effectuées une à une

- a/ MeSH : Surgery
 LCSH : Surgery
 Surgeons
 Surgery, Experimental
- b/ MeSH : Bone and Bones
 Densitometry
 LCSH : Bone densitometry
- c/ MeSH : Inservice Training

LCSH : subdivision in- service training under occupational groups and types of employees

Initialement 9824 vedettes LCSH étaient liées à 8799 vedettes MeSH et 9823 vedettes MeSH sont liées à 9427 vedettes LCSH. Dans la deuxième phase de ce projet la liaison des données sont entrées dans les listes d'autorités de MeSH et LCSH. Lorsque cette phase du projet sera mis à exécution, ce ne sont pas les champs d'entrées au sujet qui définissent le format d'autorité MARC. En conséquence les données affichées sont entées dans un champs d'affichage du sujet localement défini.

D'après Olson, la plus importante des tâches est d'utiliser les résultats pour résoudre le problème

(3) en prévoyant des liaisons entre les vedettes correspondantes dans les systèmes associés. Une fois que les vedettes affichées sont dans les listes d'autorité MARC qui peuvent être indexées . Il serait donc possible aux concepteurs de catalogues de bibliothèques automatisées d'ajouter à leurs systèmes la capacité d'indexer ces listes.

4. 8 Identification des systèmes de recherche :

Catalogues en ligne utilisant plusieurs vocabulaires contrôlés

Nom du système	Type de vocabulaires contrôlés	localisation	domaine couvert
CABI : CAB International	Macrovocabulaire : universal agricultural thesaurus unifying		Agriculture
DOBIS : Dortmunder Bibliotheks System	Liste alphabétique intégrée de vedettes matières : LCSH, CSH, CUT	University of Dortmund- Allemagne developpé par National Library of Canada	Multidisciplinaire
ETHICS	Classification universelle : CDU, liste de vedettes matières	Institut Fédéral de Technologie de Zurich (Suisse)	Sciences et techniques
FAO : food agriculture Organization of the United Unions	Macrovocabulaire : Universal Agricultural Thesaurus Unifying		agriculture

INTERCONCEPT	Banque de descripteurs	Unesco	Sciences Sociales
Library of Congress	Liste de vedette matière LCSH comme superstructure + microthesauri : LIV, 4 TGM, AAT*, SHCL	Bibliothèque du Congrès. Washington	Multidisciplinaire
NAL : National Agriculture Library	Macrovoculaire : Universal agricultural thesaurus	USA	Agriculture
NOTIS : Northwestern Online Total Integrated System	Liste de vedettes matières : LSH, MeSH	USA	Sciences médicales
ORION : Online Integrated Library System	2 Listes de vedettes matières : MeSH et LCSH	University of California, Los Angeles Library	multidisciplinaire
PRECIS : Preserved Context System	Vocabulaire intermédiaire	British National Library	multidisciplinaire
ROOT	Microthesauri	British Standard Institution (BSI)	4
SCORPIO / NELVYL	Liste de vedettes matières LCSH	Bibliothèques universitaires (USA)	
UMLS : Unified Medical Language System	Vocabulaires fusionnés produisant une liste alphabétique unifiée : Metathesaurus basé sur les listes de vedettes matières : MeSH, LCSH, SNOMED, CMIT, PDQ *	National Library of Medicine	Sciences médicales
UTLAS : University of Toronto Library Automation System		USA	
WLN : Western Library Network		Canada, USA	

Bases de données utilisant plusieurs vocabulaires contrôlés

BRS/ TERM	index matières : ERIC, MEDLINE, PSYCINFO, Sociological Abstracts, texte libre	USA	
CEE : Communauté économique européenne	SNOMED, ICD, EMTREE	CEE	
DIALOG Informatio n Retrieval service	DIALINDEX	USA	tout les domaines confondus
LIV : List of integrate vocabularies	DIALOG, Medline, ORBIT		

5. Conclusion

Si la tendance actuel offre langages naturels et langages contrôlés ou beaucoup plus le langage naturel que les vocabulaires contrôlés cela est dut au fait que les vocabulaires contrôlés n'offre pas des résultats toujours des satisfaisants. L'une des raisons est que ce type de langages n'est toujours accessibles à un large public d'utilisateurs. Néanmoins, ces derniers s'avèrent un outil documentaire évident pour la recherche dans les domaines de connaissances spécialisées, puisqu'ils offrent un accès précis au sujet et plus de rappels possible. Nous pouvons conclure qu'au niveau de la recherche au sujet, les vocabulaires contrôlés servent surtout à compenser les insuffisances du texte libre dans des systèmes regroupant des catalogues de plusieurs disciplines appartenant à un domaine plus général.

3. Bibliographie

3.1 Généralités

[1] **ALBRECHTSEN, H., OERNAGERS, S.** Knowledge organization and quality management. In Proceedings of the third ISKO conference. Frankfurt : Indeks Verlag, 1994. 457p.

[2] **American Psychological Association.** Thesaurus of psychological index terms. 6ème éd. Arlington, Virginia : Author, 1991

[3] **BEKIN, N. J., CROFT, W.B.** Retrieval techniques. *Annual Review of Information Science and Technology*, 1987, vol. 22, p. 109-145

[4] **BLAIR, D.C.** Language and representation in information retrieval. New York : Elsevier, 1990

[5] **BROADBENT, E.** Classification access in the online catalog. *Cataloging and Classification Quarterly*. 1995, vol. 21(2), p. 119-142

[6] **BYRNE, Alex, MICCO, M.** Improving OPAC subject access : the ADFA experiment. *College and Research Libraries*, sept. 1988, vol. 49, p. 432- 441

[7] **CHADWICK, T. B.** Searching across files and systems : multifile search techniques. Wilton, Connecticut : Oline In, 1992. P. 32- 36

[8] **CHAMIS, Alice Y. et al.** Effectiveness and efficiency of searchers in online searching : preliminary results from a study of information seeking and retrieving. In Proceedings of the American Society for Information Science 49 th Annual Meeting, 1986, vol. 23

[9] **CHAMIS, A. Y.** Online database search strategies and the thesaural relationship models. PHD Thèse de Doct. : Cleveland , Western Reserve University, Ohio, 1984

[10] **CHERRY, Joan M.** Improving subject access in OPACs : an exploratory study of conversion of users' queries. *Journal of Academic Librarianship*, may 1992, vol. 18 (2), p. 95-99

[11] **CHOWDHURY, G. G., NEELAMEGHAN, A., CHOWDHURY, S. VOCON :** Vocabulary Control online in Miro-ISIS databases. *Knowledge Organization*. 1995, vol. 22, p. 18- 22

[12] **COCHRANE, Pauline A., MARKEY, Karen.** Catalog use studies - since the introduction of the online interactive catalogs : impact on design for subject access. *Library and Information Science Research*. 1983, vol. 5, p. 337- 363

- [13] **COCHRANE, Pauline A.** access in the online catalog. *Research Libraries in OCLC : a Quarterly*. janv. 1982, n° 5, p. 1-7
- [14] **FURNISS, Peter.** A study of the compatibility of two subject catalogues. These : Sheffield University, Postgraduate School of Librarianship and Information Science (UK), 1980
- [15] **GILES-PETERS, A.** Experiments in the mechanical construction of cross- database thesauri. In Raitt David I., Jeapes, Ben Proceedings of the 17 th International Online Information Meeting. London, december 7-9 1993. Oxford, new Jersey : Learned Information (Europe), 1993. p. 137- 147
- [16] **GÔDERT, W., SILK, H.** The Design of subject elements in Online Public Access Catalogs. *International classification*, 1990, vol. 17 (2), p. 66- 76
- [17] **HANCOCK- BEAULIEU, M.** Evaluating the impact of an online library catalogue on subject searching behaviour at the catalogue and the shelves. *Journal of Documentation*. 1990, vol.46 (3), p. 318- 338
- [18] **HILDRETH, C. R.** Online browsing support capability. In Petrarca, Anthony E., Taylor, Celianne I. et Kohn, Roberts. *Information Interaction : Proceedings of the 45 th ASIS Annual Meeting*. Columbus, Ohio ; New York, Plains : Knowledge Industry Publications, 19..., p. 127-132
- [19] **HILDRETH, C. R.** Online public access catalogs. In Williams Martha E. *Annual Review of Information Science and Technology*, vol.20, 1985, p. 233- 285
- [20] **International Organization for Standardization.** ISO 2788 : Guidelines for the establishment and development of monolingual thesauri.- 2 ème éd.- Geneva : ISO, 1986
- [21] **KRESLINS, K., O'BRIEN, A., SMITH, I. A.** Online public access catalogues in Latvia : strategies for subject access
- [22] **MICCO, M.** The Next generation of online public access catalogs : a new look at subject access using hypermedia. *Cataloging and Classification Quarterly*. 1991, vol. 13, p. 103- 132
- [23] **NEVILLE, H. H.** Feasibility study of a scheme for reconciling thesauri covering a common subject. *Journal of Documentation*. 1970, vol. 26 (4), p. 313- 336
- [24] **NOHR, Holger.** Subject access via keywords in on- line public catalogues (OPACS) *Libri*. 1991, vol. 41 (3), p. 165- 169
- [25] **ROLLAND- THOMAS, P.** Preliminary study of the development of controlled vocabulary in french for public and school library. *Documentation et Bibliothèques*, janv.-mars 1993, vol. 39, p. 11- 18
- [26] **ROWLEY, J.** The Controlled versus natural indexing languages debate revisited : a perspective on information retrieval practice and research. *Journal of Information Science*. 1994, vol. 20 (2), p. 108- 119

- [27] **SCIBOR, E., JABRZEMSKA, E.** Thesaurus of Common Topics. In Kent, A. *Encyclopedia of Library and Information Science*, vol. 44. New York, 1989. p.388- 395
- [28] **SOERGEL, D.** Indexing languages and thesauri : construction and maintenance. Los Angeles : Melville Pub., 1974
- [29] **SVENONIUS, Elaine.** Compatibility of retrieval languages : introduction to a forum. *International Classification*, 1983, vol. 10 (1), p. 2- 4
- [30] **UNESCO.** Compatibilité between indexing and retrieval languages, Ohio, 1982
- [31] **UNESCO.** Unified system of information retrieval languages. Latvia, 1977
- [32] **UNESCO.** Overcoming the language barrier, Luxembourg, 1977
- [33] **UNESCO.** UNISIST : study report on the feasibility of a world science information system. Paris : Unesco Press, 1971
- [34] **VITUKHNOVSKAYA, A. A.** On the theory of compatibility of information retrieval thesauri. *Nauchnye- i Tekhnicheskie Biblioteki SSSR*, 1976, n° 11, p. 6- 14
- [35] **WINKE, R. Conrad.** Discarding the main entry in an online cataloging environment. *Cataloging and Classification Quarterly*. juin 1993, vol. 16 (1), p. 50- 70
- [36] **YASAR, Tonta.** Analysis of search failures in document retrieval systems : a review. *The Public- Access Computer Systems Review*. 1992, vol. 3 (1), p. 4- 53
- [37] **YOSHIDA, S.** Subject access in the online age : reviews of classification and indexing literature. *Toshokan- Kai (The Library world)*. mai 1993, vol. 45 (1), p. 143- 150

3.2 Intégration et fusion de Vocabulaires

- [38] **BELING, G., WERSIG, G.** The New Concept of an Intermediary Languages System for Information Networks. In Batten, W. E. *EURIM II : A European Conference on the Application of Research in Information Services and Libraries*. London : Aslib, 1977. P.117- 121
- [39] **BIRU, T., EL- HAMDouchi, A., REES, R.S.** Inclusion of relevance information in the term discrimination model. *Journal of Documentation*, 1989, vol. 45, p. 85-109
- [40] **CAMMELI, A.** THES GIUR : a thesaurus as an aid to a legal data base and for the analysis of the lexicon of the law. In Negrini, Giliola, Farnesi, Tamara et al. *Documentary languages and databases Conference*, Rome, 3- 4 december, 1990
Advances in Knowledge Organization, 1991, p. 85- 101
- [41] **CHAMIS, Alice Y.** Selection of online databases using switching vocabularies. *Journal of the American Society for Information Science*. 1988, vol. 39, p. 217- 218

- [42] **CHAMIS, A. Y.** Research on Information seeking and Retrieving : a progress report. In Proceedings of the National Online Meeting, 7 th Annual Meeting, 1986. p. 67- 72
- [43] **CHAMIS, A. Y.** The Usefulness of switching vocabularies for online databases. *In ASIS 85 : Proceedings of the 48 th ASIS Annual Meeting* 1985. Las Vegas, 20- 24 oct. 1985, vol. 22, p. 311- 314
- [44] **CHAMIS, A. Y.** Vocabulary control and search strategies in online searching. Westport : Greenwood Press, 1991. 121 p.
- [45] **CHAN, Lois Mai.** Functions of a subject authority file. In Markey Drabenstott, Karen *Subject authorities in the online environment*. Chicago, London : American Library Association, 1991. p. 9- 30
- [46] **CHAN, Lois Mai.** Alternatives in Library of Congress Classification. *Cataloging & Classification Quarterly*, 1995, vol. 19 (3- 4), p. 67- 87
- [47] **CHAN, L. M.** Functions of a Subject Authority File. In Markey Drabenstott, Karen *Subject Authorities in th online Environment*. Chicago, London : American Library Association, 1991. P. 9- 30
- [48] **CHAN, L. M.** (Page consultée le 04 mars 1997). Lois Mai Chan Professor, [en ligne]. Adresse URL : [http : // www. Uky. Edu...ies / SLIS / 1 chan. Html](http://www.Uky.Edu...ies/SLIS/1chan.Html)
- [49] **CHAN, L. M.** Subject analysis tools online : the chalenge Ahead. *Information Technology and Libraries*, sept. 1990, vol.9 (3), p. 258- 262
- [50] **CHAN, L. M., POLLARD, R.** Thesauri used in online databases : an analytical guide. NewYork : Greenwood Press, 1988
- [51] **CHAN, L. M.** The Library of Congress Classification System in an online Environment. *Cataloging & Classification Quarterly*, 1990, vol. 11 (1), p. 7- 25
- [52] **CHAN, L. M., POLLARD, R.** Thesauriused in Online Dtatbase : an analytical guide. Wesport : Greewood Press, 1988
- [53] **CHAPLAN, M. A.** Mapping laborline thesaurus terms to Library of Congress Subject Headings : implications for vocabulary switching. *Library Quarterly*, janv. 1995, vol. 65, n° 1, p. 39-61
- [54] **COATES, E. J.** Switching languages for indexing. *Journal of Documentation*, 1970, vol. 26 (2), p. 102- 110
- [55] **CONGREVE, J.** Problems of subject access : automatic generation of printed indexes and online thesaural control. *Program*, 1986, vol. 20 (2), p. 204- 210
- [56] **CONNELL, Tsherra H.** Techniques to improve subject retrieval in online catalog. *Information Technology and Libraries*, june 1992, vol. 10 (2), p. 87- 98

- [57] **COUSINS, S. A.** Enhancing subject access to Opacs controlled vocabulary vs natural language. *Journal of Documentation*, 1992, vol. 48 (3), p. 291- 309
- [58] **COUSINS, S. A.** In their own words : an examination of catalogue users' subject queries. *Journal of Information Science*, 1992, vol. 18, p. 329- 341
- [59] **CRAWFORD, W.** The future online catalog : a single view of multiple database. *Information Technology and Libraries*, june 1993, vol. 12 (2), p. 253- 254
- [60] **FROST, Carolyn O.** Subject heading compatibility between LCSH and catalog files of a large research library : a suggested model for analysis. *Information Technology and Libraries*, 1988, vol. 7 (3), p. 288- 299
- [61] **FIELD, B. J.** A Thesaurus based indexing and classification system developed for INSPEC products and services. *Journal of Documentation*, mars 1974, vol. 30 (1), p. 1-17
- [62] **FIELD, B. J.** Towards automatic indexing : relationship between free and controlled languages indexing and automatic generation of controlled subject headings and classification. INSPEC, 19 ?
- [63] **FORSYTH, R., RADA, R.** *Machine learning : Applications in expert systems and information retrieval*. Chicester, UK : Ellis Horwood, 1986
- [64] **DAHLBERG, I.** Conceptuel compatibility of ordering systems. *International Classification*. 1983, vol. 10 (1), p.5-8
- [65] **DAHLBERG, I.** The Broad system for ordering (B S O) as a basis for an integrated social sciences thesurus. *International Classification*. 1980, vol. 7 (2), p. 66-72
- [66] **DAHLBERG, I.** Towards establishment of compatibility between indexing languages. *International Classification*, 1981, vol. 8 (2), p. 86- 91
- [67] **DEXTRE, S. G., Clarke, T. M.** A System for machine-Aided thesaurus construction. *ASLIB Proceedings*, 1981, vol. 33 (3), p. 102-112
- [68] **DOSZKOCS, TE., SASS, R. K.** An Associative semantic network for machine-aided indexing classification and searching. In Fidel, R., Kwasnik, B. H., Smith, P. J. *Proceedings of the 3 rd ASIS SIG/ G Classification Research Workshop*, 1989
- [69] **DUBOIS, C. P. R.** Free text vs : controlled vocabulary a reassessment. *Online Review*, 1987, vol. 11 (4), p. 243-
- [70] **DYKSTRA, M.** LC subject headings disguised as a thesaurus. *Library Journal*, mars 1988, p. 42- 46
- [71] **EIN-DOR, P., SPIEGLER, I.** Naturel language access to multiple databases : a model and a prototype. *Journal of management Information Systems*. 1995, vol. 12 (1), p. 171- 197

[72] **FENG, Suliang**. A comparative study of indexing languages in single and multidatabase searching : Library and Information Science Abstracts, Information Science Abstracts, ERIC (Educational Resources Information Center). *Canadian Journal of Information Science*, mai 1989, vol. 14 (2), p. 26- 46

[73] **FIDEL, Raya**. Moves in Online Searching. Online Review, febr. 1985, vol. 9, p. 61- 76

[74] **FIDEL, Raya**. Online searching styles : a Case- Study Stacks Based Model of Searching Behaviour. *Journal of ASIS*, july 1984, vol. 35, p. 211- 221

[75] **FIDEL, Raya**. Searchers' selection of search keys : I. The Selection routine. II. Controlled vocabulary or free- text searching. III. Searching styles. *Journal of the American Society for Information Science (ASIS)*, august 1991, vol. 42 (7), p. 490- 527
II. Controlled vocabulary or free- text searching. p. 501- 514

[76] **FIDEL, Raya**. Terminological Knowledge structure for intermediary expert system. *Information Proceesing and Management*, janv.- fevr. 1995, vol. 31 SHATFORD (1), p. 15- 27

[77] **FIDEL, R**. Who needs controlled vocabulary ? *SpecialLibraries*. 1992, vol. 83, p. 1- 9
GARLAND, K. An experiment in automatic hierarchic document classification. *Information Processing and Management*, 1983, vol. 19, p. 113- 120

[78] **HOOD, M. W., EBERMANN, C**. Reconciling the CAB thesaurus and AGROVOC. *Quarterly Bulletin of the International Association of Agricultural Librarians and Documentalists*, 1990, vol. 35, p. 181- 185

[79] **HORSNELL, V**. Evaluation and implementation of an intermediate lexicon. In Batten, W. E. *EURIM II Proceedings*. London : ASLIB, 1977. P. 110- 112

/80/ **HOWARTH, L. C**. Online access to classification numbers : quality control in canadian bibliographic database. *Library Hitech*. 1994, vol. 12 (4), p. 71- 81

[81] **HUG, H., WALSER**. Retrieval in the ETH database using the UDC. In Fugmann, R . Tools for knowledge organization and the Human Interface, august 14- 17. Frankfurt : Indeks Verlag, 1991. Vol. 2, p. 202- 209

[82] **HUMPHRIES, A. W., NAISWALD, G. V**. Developing a quality assurance program for online services. *Bulletin of the Medical Library Association*. juillet 1991, vol. 79 (3), p. 263- 270

[83] **HUSAIN, S., O'BRIEN, A**. Recent trends in sibject access to OPACs : an evaluation. *International Classification*. 1992, vol. 19 (3), p. 140- 145

[84] **ITO, H., FUKUMURA, T**. Financial advisor : consultation system for financing information retrieval by multi-dimensional thesaurus. In *1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent systems for the 21 th century*, Vancouver, oct. 22- 25, 1995. New York : IEEE., 1995, vol. 2, p. 1160- 1165

- [85] **KLEINBART, P.** Prolegomenon to intelligent thesaurus software. *Journal of Information Science*, 1985, vol. 11, p. 45- 53
- [86] **KNAPP, Sara D.** Creating BRS / TERM, a vocabulary database for searchers. *Database*, 1984, vol. 7, p. 70- 75
- [87] **LANCASTER, F. W.** Compatibility and convertibility. In Vocabulary control for information retrieval. 2ème éd. Arlington, Virginia : Information Resources Press, 1986. chap. 19, p.270-
- [88] **LANCASTER, F. W.** Evaluation of interactive knowledge based systems : overview and design for empirical testing. *Journal of the American Society for Information Science*, janv. 1996, vol. 47 (1), p. 57- 69
- [89] **LANCASTER, F. W., ELLIKER, C., CONNELL, T. H.** Subject analysis. In Williams, M. E. *Annual Review of Information Science and Technology*. Washington : American Society for Information Science, 1989.- vol. 24, p.35- 84
- [90] **LANCASTER, F. W., SMITH, Linda C.** Compatibility issues affecting information systems and services. Paris : Unesco General Information Programme and UNISIST, 1983. p. 50- 77
- [91] **LARSON, Ray R.** The Decline of subject searching in online catalogs. 1989
- [92] **LARSON, Ray R.** Managing information overload in online catalog subject searching. In *Managing Information and Technology, Proceeding of the 52 nd Annual Meeting of the american society for information science ASIS*, Washington, 30 oct- 2 nov. 1989. Washington : ASIS, 1989. vol. 26
- [93] **LEVY, F.** Compatibility between classification and thesauri : evaluation of a first study in the of information storage and retrieval : paper presented at the conference of the international federation for documentation. Tokyo, 1967
- [94] **MANDEL, Carol A.** Cataloguing for access. *Reference Librarian*. 1991, n° 34, p. 61- 68
- [95] **MANDEL, C. A.** Enriching the library catalog record for subject access. *Library Resources and Technical Services*. 1985, vol. 29, p. 5- 15
- [96] **MANDEL, C. A.** Multiple thesauri in online library bibliography systems : a report prepared for Library of Congress Processing services. Washington : Library of Congress, Cataloging Distribution Service, 1987. 94 p.
- [97] **MANDEL, C. A., HERSCHMAN, Judith.** Online subject access-enhancing the library catalog. *Journal of Academic Librarianship*. juillet 1983, vol. 9 (3), p. 148- 155
- [98] **MANIEZ, Jacques.** Les Langages documentaires et classificatoires : conception, construction et utilisation dans les systèmes documentaires. Paris : Editions d'Organisation, 1987

- [99] **MANIEZ, J.** Relationships in thesauri : some critical remarks. *International Classification*, 1988, vol. 15, p. 133- 138
- [100] **MARKEY, Karen.** Integration the machine- readable LCSH into online catalogs. *Information Technology and Libraries*. sept. 1988, vol. 7 (3), p. 297-312
- [101] **MARKEY, Karen.** Subject- searching experiences and needs of online catalog users : implications for library classification. *Library Resources and Technical Services*. 1985, vol. 29, p. 34- 51
- [102] **MARKEY, K.** Subject searching in library catalogs : before and after the introduction of online catalogs. Dublin, Ohio : OCLC, Inc, 1984
- [103] **MARKEY, Karen.** Searching and browsing the Dewey Classification in an online catalog. *Cataloging and Classification Quarterly*. 1987, vol. 7 (3), p. 37- 68
- [104] **National Information Standards Organization.** Z 39. 19-199x : *Standard guidelines for the construction, format and maintenance of monolingual thesauri* . Washington , DC : NISO, 1990
- [105] **NEELAMEGHAN, A.** Absolute syntax and structure of an indexing and switching language. In Neelameghan, A. *Ordering systems for global information network*. Bangalore, India : International Federation for Documentaion. 1979, p. 165- 176
- [107] **NIEHOFF, R. T.** Development of an integrated energy vocabulary and the possibilities for on- line subject switching. *Journal of the American Society for Information Science*. 1976, vol. 27, p. 3- 17
- [108] **NIEHOFF, R. T.** The Optimization and use of automated subject switching for better retrieval. In *Proceedings of the American Society for Information Science 43 rd Annual Meeting*. 1980, vol. 17
- [109] **NIEHOFF, Robert T. et al.** The Design and evaluation of a vocabulary switching system for use in multi-base search environments. Columbus : Battelle, 1980
- [110] **NIEHOFF, R. T., KWASNY, S.** The Role of automated subject switching in a distributed information network. In *Second International Online Information Meeting*. New York : Learned Information, 1978. *Online Review*. 1979, vol. 3, p. 181- 194
- [111] **NIEHOFF, R. T., KWASNY, S., WESSELLS, M.** Overcoming the database vocabualry barrier : a solution. Online, 1979, vol. 3 ; p. 46- 54
- [112] **NIEHOFF, R. T. MACK, G.** Final report on evaluation of the vocabulary switching system. Columbus : Battelle, 1984
- [113] **NIEHOFF, R. T., MACK, G.** Evaluation of the vocabulary switching system : final report. Columbus, Ohio : Battelle Columbus Laboratories, 1984

- [114] NIEHOFF, R. T., MACK, G. The Vocabulary switching system : description of evaluation studies. *International Classification*. 1985, vol. 12, p. 2- 6
- [115] NOMOTO, K., KUBOT, T., KOSUGE, Y. Fuzzy thesaurus generation based on crass- index matrix for case- based reasoning. In IEEE International conference on systes, Man and Cybernetics..., Vancouver, oct. 22- 25 1995. New- York : IEEE, 1995, vol.5 , p. 4033- 4038
- [116] NYE, Julie Blume. User interaction with the authority structure of the online catalog : results of a survey. *Information Technology and Libraries*. 1988, vol. 7 (3), p. 313- 316
- [117] O'BRIEN, Ann. (Page consultée le 27 mars 1997). Online catalogs : enhancements and developments, [en ligne]. adresse URL : [http : // www. Asis. or...t-94-chapter6. html](http://www.Asis.or...t-94-chapter6.html). In Williams, Martha E. *Annual Review of Information Science and Technology*, vol. 29. Medford, New Jersey : ASIS, 1994. p. 219- 242
- [118] O'BRIEN, Ann. Relevances as a measures of evaluation in OPACs. *Journal of Information Science*, august 1990, vol. 16 (4), p. 265- 271
- [119] O'LEARY, M. Databases of the 90's : the age of access. *Database*, 1990, vol. 13 (2), p.15- 21
- [120] PEJTERSEN, A. M. A New approach to design of document retrieval and indexing systems for OPAC user. In Raitt, D.I., Jeapes, B. *Online Information 93. 17 th International Online Information Meeting Proceedings*. Oxford : Learned Inf., 1993. p. 273- 290
- [121] PITERNICK, Anne B. Decision Factors Favouring the Use of Online Sources for Providing Information. *R Q*, summer 1990, vol. 29, p. 534- 544
- [122] PITERNICK, A. B. Functions and capabilities of Online Search Systems : A Checklist. *Online Review*, dec. 1989, vol. 13, p. 457- 476
- [123] PITERNICK, A. B. Vocabularies for online subject searching. In Kent, A. *Encyclopedia of Library and Information Science*, vol. 45 . New York : Marcel Dekker, 1990. p. 399- 420
- [124] PETERSON, Tony. Multiple authorities in library system. Unpublished paper. In Art Libraries Society / North America Symposium on Authority Control, febr. 1986
- [125] Knowledge Sources : Multiple Thesauri. (Page consultée le 19 mars 1997), [en ligne]. Adresse URL : [http : // ai. Bpa. Arizona...bsection 3-6-1. Html](http://ai.Bpa.Arizona...bsection3-6-1.html)
- [126] RADA, R. Connecting thesauri : issues and cases. *International Classification*, 1987, vol. 14, p. 63- 69
- [127] RADA, R. Maintaining thesauri and metathesauri. *International Classification*, 1990, vol. 17, p. 158- 164
- [128] RADA, R., BARLOW, J., POTHARST, J. et al. Document ranking using enriched thesaurus. *Journal of Documentation*, 1991, vol. 47, p. 227- 332

- [129] **RADA, R., BICKNELL, E.** Ranking documents with a thesaurus. *Journal of the American Society for Information Science*, 1989, vol 40, p. 304- 310
- [130] **RADA, R. DARDEN, L. et ENG, J.** Relating two knowledge bases : the rôle of identity and part-whole. In Boudreaux, T. C., Hamill, B. W. *The Rôle of language in problem solving 2*. Amsterdam : North- Holland, 1987.
- [131] **RADA, R., MARTIN, B. K.** Augmenting thesauri for information systems. *ACM Transactions on Office Information Systems*, 1987, vol. 5, p.378- 392
- [132] **SAGER, J. C. et al.** Thesaurus integration in the social sciences. Part II Stages towards integration ; Part III Guildelines for the integration of thesauri. *International Classification*. 1982, vol. 9 (1), p. 19- 26 ; vol. 9 (2), p. 64- 70
- [133] **SCHATZ, B. R., JOHNSON, E. H., COCHRANE, P. A. et CHAN, H.** Interactive term suggestion for users of digital libraries : using subject thesauri and co-occurrence lists for information retrieval. In proceedings ACM Digital Libraries, Bethesda, MD, 1996
- [134] **STAM, Deidre.** Choosing our words : Reflections on Authority Control in Art Information Systems. Inpublished paper, presented at Art Libraries Society / North America Symposium on Authority Control, february 1986
- [135] **SILVER, J. P., KILINGBIEL, P. H.** An Operational system for subject switching between controlled vocabularies. *Information Processing and Management*. 1993, vol. 29, p. 47- 59
- [136] **SMITH, P. J. KRAWCZAK, D., SHUTE, S. et al.** Bibliographic information retrieval systems : using al techniques to improve cognitive compatibilité. In *Proceedings of the 1986 IEEE Conference on Systems, Man and cybernetics*, 1986. p. 395-
- [137] **SNEIDERMAN, C. A., BICKNELLI, E. J.** Computer- assisted dynamic integration of multiple medical thesauri. *Computers in Biology and Medicine*. janv.- mars 1992, vol. 22 (1- 2), p. 133- 145
- [138] **SOERGEL, D.** A General model for indexing languages : the basis for compatibility and integration. In Wellisch, H. et Wilson, T. D. *Subject retrieval in the seventies : new directions*. Westport : Greenwood, 1972. p. 36- 61
- [139] **SU, L. T.** Evaluation measures for interactive information retrieval. *Information Processing & Management*, 1992, vol. 28 (4), p. 503- 516
- [140] **SVENONIUS, Elaine.** Design of controlled vocabularies... In Kent A. *Encyclopedia of library and information science*. New York ; Basel ; Hong Kong : Marcel Dekker. 1990, vol. 45, p. 82- 109
- [141] **SVENONIUS, Elaine.** Unanswered Question in the Design of Controlled Vocabularies. *Journal of the American Society for Information Science*. 1986, vol. 37, p. 331- 340

- [142] **TAYLOR, Arlene G. et al.** Network and Vendor Authority Control. Library Resources and Technical Services, april-june 1985, vol. 29,p. 195- 205
- [143] Thesaurus of ERIC descriptors. 12 th éd. Phoenix, Arizona : Oryx Press, 1990
- [144] **VASILJEV, A.** Enhancement of the subject access vocabulary in an online catalogue. . In Fugmann, R.*Tools for Knowledge Organization and the Human Interface*, Proceedings 1st International ISKO- Conference.- Frankfurt, Main : Indeks Verlag. 1990, p. 163- 170
- [145] **VIZINE-GOETZ, D., MARKEY, K.** Subject authority control in online catalog design. In OPACs and Beyond. Dublin : OCLC Computer Library Center, 1989. p. 41- 50
- [146] **WEINBERG, B. H.** The Hidden classification in Library of Congress Subject Headings for Judaica. *Library Resources and Technical services*. oct. 1993, vol. 37 (4), p. 369- 379
- [147] **WEINBERG, B. H.** Library classification and information retrieval thesauri : comparison and contrast. *Cataloging and Classification Quarterly*. 1995, vol. 19 (3/4), p. 23- 44
- [148] **WEINBERG, B. H.** Why post coordination fails the searcher. *Indexer*. avril 1995, vol. 19 (3), p. 155- 159
- [149] **WEINTRAUB, Tamara S.** The Dual- thesaurus catalog : MeSH and LCSH. *Technicalities*, dec. 1992, vol. 12 (2)
- [150] **WEINTRAUB, I.** The Terminology of alternative agriculture searching AGRICOLA, CAB and AGRIS. *Quarterly Bulletin of the International Association of agricultural librarians and documentalists*, 1992, vol. 37 (4), p. 209- 213
- [151] **WESSELS, Michael B., NIEHOFF, R. T.** Synonym switching and authority control. In Ghikas, Mary W. *Authority control : the key to tomorrow's catalog*. Phoenix : Oryx Press, 1982, p. 97- 118
- [152] **YEE, Marthe M., , Layne S.** Online public access catalogs. In Kent A. *Encyclopedia of Library and Information Science*. 1991 , vol. 35, p. 149- 230
- [153] **YEE, M. M.** System design and cataloging meet the user : user interfaces to Online Public Access Catalogs. *Journal of ASIS*, 1991, vol. 42
- [154] **YEE, M. M., SOTO, R.** User problems with access to fictional characters and personal names in online access catalogs. *Information and technology libraries*, 1991, vol. 10, p. 3- 13
- [155] **ZENG, Lei.** Achieving compatibility of indexing languages in online access environment : online subject searching in a multithesauri environment. In Kent A. *Encyclopedia of Library and Information Science*. 1992, vol. 50 (suppl. 13). p. 1-24
- [156] **ZENG, Lei.** Establishing a compatible general vocabulary in China : the capability. *International Classification*, 1990, vol. 17 (2), p. 91- 98

[157] **ZENG, Lei**. Research and development of classification and thesauri in China. In *5th International Study Conference on Classification Research*, Toronto, Canada, June 24- 28, 1991

3.2.1 Classification clustering

[158] **ACM**. ACM Computing reviews classification system. *ACM Computing Review*, 1994, vol. 35, p. 4- 44

[159] **ALLEN, R. B., OBRY, P., LITTMAN, M.** An interface for navigating clustered document sets returned by queries. in proceedings ACM SIGOIS : Conference on Organizational Computing Systems (COOCS), Milpitas, nov. 1993. p. 166- 171

[160] **ALLEN, R. B.** (Page consultée le 02 avril 1997). Retrieval from facet spaces, [en ligne]. Adresse URL : <http://superbook.b...FACETS/facets.html>

[161] **ALLEN, R. B.** Two digital library interfaces which exploit hierarchical structure. In proceedings of Electronic Publishing and the Information Superhighway, Boston, May 1995. p. 134- 141

[162] **CROFT, W. B.** Clustering large files of documents using the single- link method. *Journal of the American Society for Information Science*. 1977, vol. 28, p. 341- 344

[163] **CROFT, W. B.** A Model of cluster searching based on classification. *Information Systems*, 1980, vol. 5, p. 189- 195

[164] **CROUCH, C. J.** An analysis of approximate vs exact discrimination values. *Information Processing and Management*. 1988, vol. 24, p. 5- 16

[165] **CROUCH, C. J.** Cluster- based approach to thesauri construction. in *Proceedings of the 11th International Conference on Research and Development in Information Retrieval. Grenoble* : Presses Universitaires de Grenoble, 1988

[166] **GRIFFITHS, A., LUCKHURST, H. C., et WILLETT, P.** Using inter-document similarity information in document retrieval systems. *Journal of the American Society of Information Science*. 1986, vol. 37, p. 3- 11

[167] **GRIFFITHS, A., ROBINSON, L. A. et WILLETT, P.** Hierarchic agglomerative clustering methods for automatic document classification. *Journal of Documentation*. 1984, vol. 40, p. 175- 205

[168] **HUESTIS, Jeffrey C.** Clustering LC classification numbers in an online catalog for improved browsability. *Information Technology and Libraries*, vol. 7 (4)p. 381- 393

[169] **JARDINE, N., Van RIJSBERGEN, C. J.** Use of hierarchic clustering in information retrieval. *Information Storage and Retrieval*. 1971, vol. 7, p. 217- 240

- [170] **LARSON, R. R.** Classification, clustering, probabilistic information retrieval and the online catalog. *Library Quarterly*, 1991, vol. 6, p. 133- 173
- [171] **LARSON, R. R.** Evaluation of advanced retrieval techniques in an experimental online catalog. *Journal of the American Society for Information Science*, 1992, vol. 43, p. 34- 53
- [172] **LARSON, R. R.** Experiments in automatic Library of Congress Classification. *Journal of the American Society for Information Science*, 1992, vol. 43, p. 130- 148
- [173] **O' KANE, K. C.** A design for a world wide web Document Retrieval System. *Online and CD ROM Review*. 1996, vol. 20, n° 1, p. 11- 20
- [174] **O'KANE, K. C.** A language for implementing information retrieval software. *Online Review*, 1992, vol. 16 (3), p. 127- 137
- [175] **O'KANE, K. C.** An information retrieval implementation language. ACM Mid-Southeast Chapter Fall Conference, 6- 7 nov. 1991
- [176] **O'KANE, K. C.** Design of a text- based information storage and retrieval system in MUMPS. *MUG Quarterly*. 1991, vol. 21 (4), p. 21- 26
- [177] **O' KANE, K. C.** Generating document indices and clusters from common denominators in large document collections. *Information Processing and Management*. 1996, vol. 32 (1), p. 105- 115
- [178] **O' KANE, K. C.** Language independent generation of hierarchical document indices in large documents collections. Institut fuer Analytische Chemie, Technische Universitaet Dresden, mai 1995. p. 11- 12
- [179] **O'KANE, K. C.** (page consultée le 04 février 1997). Kevin Charles O'Kane, [en ligne]. adresse URL : [http : // www. cs. Uni. edu/ ~Okane/ resume. html](http://www.cs.Uni.edu/~Okane/resume.html)
- [180] **O'KANE, K. C., HILDEBRANDT, R. J.** AN integrated health information processing and retrieval system. In *Proceedings AFIPS NCC & E*, 1974, vol. 43, p. 93- 105
- [181] **O'KANE, K. C., Mc COLLIGAN, E. E.** A ; object oriented class library for medical software development. *Methods of Information in Medicine*, 1996, vol. 34 (4)
- [182] **SALTON, G.** Another look at automatic text retrieval systems. *Communications of the ACM*, 1986, vol. 29, p. 648- 656
- [183] **SALTON, G.** *Automatic text processing*. Reading : Addison. Wesley, 1989
- [184] **SALTON, G., BUCKLEY, C.** Term- weighting approches in automatic text retrieval. *Information Processing and Management*, 1988, vol. 24, p. 513- 523
- [185] **SALTON, G., Mc GILL, M.** Introduction to modern information retrieval. New York : McGraw hill, 1983

- [186] **SALTON, G., Wong, A. et YANG, C. S.** A Vector space model for automatic indexing. *Communications of the ACM*, 1975, vol. 18, p. 613- 620
- [187] **SANJEEVI, Amba.** Computer- based linking of thesauri. Thèse de Doctorat, Ph D. : University of Alabama, 1992. Michigan : UMI Dissertation Services, 1997. 151 p.
- [188] **SANJEEVI, Amba, NARASIMHAMURTHI, N., O' KANE, K. C. et TURNER, P.** Automatic linking of thesauri. In Proceedings 19 th International ACM SIGIR : Conference on Research and Development in Information retrieval. ETH, Zurich, august 18- 22 1996
- [189] **VAN RIJSBERGEN, C. J.** Information retrieval. London : Butterworths, 1979
- [190] **VAN RIJSBERGEN, C. J., CROFT, W. B.** Document clustering : an evaluation of some experiments with the Cranfield 1400 collection. *Information Processing and Management*, 1975, vol. 11 p. 171- 182
- [191] **VOORHEES, E. M.** Implementing agglomerative hierarchic clustering algorithms for use in document retrieval. *Information Processing and Management*, 1986, vol. 22, p. 465- 476
- [192] **WILLETT, P.** Recent trends in hierarchic document clustering : a critical review. *Information Processing and Management*, 1988, vol. 22, p. 577- 597
- [193] **WILLETT, P.** Similarity coefficients and weighting functions for automatic document classification : an empirical comparison. *International Classification*, 1983, vol. 10, p. 138- 142

3.2.2 UMLS

- [194] **BARNETT, Go, CIMINO, J. J., HUPP, J. A. et al.** DX plain an evolving diagnostic decision- support system. *Journal of the American Medical Association*. juillet 1987, vol. 258 (1), p. 67- 73
- [195] **BICKNELL, E. J., SNEIDERMAN, C. A., RADA, R. F.** Computer assisted merging and mapping of medical knowledge bases. In Greenes, R. A. Proceedings of the twelfth annual symposium on computer applications in medical care, Washington, 6- 9 nov., 1988. Los Alamitos : IEEE Computer Society Press. P. 158- 164
- [196] **CHAMIS, A. Y.** (Page consultée le 04 mars 1997). Unified medical language system, [en ligne]. Adresse URL : [http : // www- camis. S...tsheet. umls95. Html](http://www-camis.S...tsheet.umls95.Html)
- [197] **CIMINO, J. J.** Representation of clinical laboratory terminology in the unified Medical Language System. In Clayton, P. D. *Assessing the value of medical informatics : proceedings fifteenth annual symposium on computer applications in medical care* (17- 20 nov. 1991, Washington . New York : Mc Graw- Hill, health professions division, 1992.- 199- 203
- [198] **CIMINO, J. J.** Use of the Unified Medical System in patient care at the Columbia- Presbyterian Medical Center. *Methods of Information in Medicine*. mars 1995, vol. 34 (1/2),

p. 158- 164. (Page consultée le 4 février 1997), [en ligne]. Adresse URL :
[http ://www.cpmc.co...pubs/ abstract. Html.](http://www.cpmc.co...pubs/abstract.html)

[199] **CIMINO, J. J.** Vocabulary and health care information technology : state of the art.
Journal of the American Society for Information Science. dec. 1995, vol 46 (10), p. 777- 782

[200] **CIMINO, J. J., AGUIRRE, A., JOHNSON, S. B.** Generic queries for meeting clinical information needs. *Bulletin of the Medical Library Association*, 1993, vol. 81 (2), p. 195- 206

[201] **CIMINO, J. J., BARNETT, Go.** Automated translation between medical terminologies using semantic definitions. In *Proceedings of the American Association for Medical Systems and Information Congress 1989*. May 10, 1989, p. 113- 117.
MD Computing, mars- avril 1990, vol. 7 (2), p. 104- 109

[202] **CIMINO, J. J., JOHNSON, S. B., PENG, P. et al.** From ICD 9- CM to MeSH using the UMLS : a how to guide. In Safran, C. In *Seventeenth annual symposium on computer applications in medical care. Ptien- Centered computing*. New York : Mc Graw- Hill, 1994. XXVIII- 984 p.

[203] **CIMINO, J. J., JOHNSON, S. B., HRIPCSAK, G., SIDELI, R. V. et al.** One year's experience with the Unified Medical Language System (UMLS) in academia and patient care. In Lun, K. C., Degoulet, P. *MEDINFO 92, Proceedings of the seventh world Congress on Medical Informatics*. Geneva Palexpo, Switzerland. sept .6- 10, 1992 (Part 2, p. 1501- 1505). Amsterdam, the Netherlands : North- Holland, Elsevier science Publishers, NTIS, 1992

[204] **CIMINO, J. J., SENGUPTA, S.** (Page consultée le 04 février 1997). IAIMS and UMLS and the Columbia- Presbyterian Medical Center, [en ligne]. Adresse URL : [http :// www. Cpmc. Co...pubs / abstract. html](http://www.cpmc.co...pubs/abstract.html)
Medical Decision Marking. 1991, vol. 11 (suppl.), s. 89- 93

[205] **CIMINO, J. J., SIDELI, R. V.** (Page consultée le 04 février 1997). Using the UMLS to Bring the Library to the Bedside, [en ligne]. adresse URL : [http :// www. cpmc. co...pubs / abstract. html](http://www.cpmc.co...pubs/abstract.html)

[206] **HUMPHREYS, B.L.** The 1994 Unified Medical Language System Knowledge Sources. *Health Libraries Review*. sept. 1994, vol. 11 (3), p. 200- 203

[207] **HUMPHREYS, B. L., LINDBERG, D. A. B.** Building the Unified Medical Language System. In Kingsland, L. C. III *Proceeding of the 3 th annual symposium of computer applications in medical care*, Washington D C, nov. 5- 8, 1989. Los Alamitos : IEEE Computer Society Press, 1989. p. 475- 480

[208] **KLATT, M. J.** An Aid for total quality searching : developing a hedge book. oct. 1994, vol. 82 (4), p. 438- 441

[209] **LINDBERG, D. A. B., HUMPHREYS, B. L. et Mc CRAY, A. T.** The Unified Medical Language System. *Methods of Information in Medicine*, 1993, vol. 32, p. 281- 291

- [210] **Mc CRAY, A. T.** The UMLS semantic network. In *Proceedings of the 13 th Annual Symposium on Computer Applications in Medical Care*, 1989. p. 503- 508
- [211] **MILLER, P. L., MORROW, J. S., POWSNER, S. M.** Semantically associated medical bibliographic retrieval : an experimental computer system. *Bulletin of Medical Library Association*, 1988, vol. p. 131- 145
- [212] **National Library of Medicine (U. S.).** (Page consultée le 29 mars 1997). Metathesaurus Data Elements and Source Vocabulary Information, [en ligne]. Adresse URL : [http : // www ; nlm. Nih...ch / umls / METAB. HTML](http://www.nlm.nih.gov/umls/METAB.HTML)
- [213] **NELSON, M. J.** Use of term co-occurrence information in information retrieval. *Canadian Journal of Information Science*, 1983, vol. 8, p. 67- 73
- [214] **OLSON, Tony.** Editorial considerations in the mapping of the MeSH and LCSH systems. In *the Medical Library Association Annual Convention*. 4 juin 1991 à San Francisco
- [215] **OLSON, Tony.** MeSH and LCSH in the Unified Subject Index. *Als Ameritech*, oct. 1995
- [216] **RICHWINE, P. W.** A Study of MeSH and UMLS for subject searching in an online catalog. *Bulletin of the Medical Library Association*, avril 1993, vol. 81 (2), p. 229- 233
- [217] **RODGERS, R. P. C.** Automated retrieval from multiple disparate information sources : the word wide web and the NLM's sourcerer project. *Journal of the American Society for Information Science*, déc. 1995, vol. 46 (10), p. 755- 764
- [218] **SHERETZ, D. D., TUTTLE, M. S., OLSON, N. et al.** Lexical mapping in the UMLS methathesaurus. In *Proceedings of the 13th Annual Symposium on Computer Applications in Medical Care*. 1989. p. 494- 499
- [219] **SQUIRES, S. J.** Access to biomedical information : the Unified Medical Language System. *Library Trends*, summer 1993, vol. 42 (1), p. 127- 151
- [220] **SPERZEL, W. D., TUTTLE, M. S.** Updating the UMLS methathesaurus : a model. In *Proceedings of the 13 th Annual Symposium on Computer Applications in Medical Care*. 1989. p. 475- 480
- [221] **SRINIVASA, P.** Optimal document- indexing vocabulary for MEDLINE. In *Information Processing and Management*. sept. 1996, vol. 32 (52). p. 503- 514
- [222] **STRAWN, Gary L.** Mapping the LCSH and MeSH. In *the Opening Session of the LITA / ALCTS Discussion Group on Authority Control in the Online Environment*. Chicago, June 24, 1990
- [223] **STRAWN, Gary L.** Multiple thesauri and Northwestern University Libraries NOTIS 5.0 Online Public Catalog. In *1992 NOTIS Users' Group Meeting*. Chicago, oct. 8, 1992

[224] **TOEWS, L.** An evaluation methodology for clinical vocabularies and evaluation of the read codes. In Olson, H. A., Wrad, D. B., *Canadian Association for Information Science, Proceedings of the 23 rd Annual Conference, Connectedness : Information, System, People , Organizations.* Edmonton, Alta (Canada) : University of Alberta, 1995. p. 71- 78

[225] **TUTTLE, M. S. SHERETZ, D . D., ERLBAUM, M. S., et al.** Implementing Meta-1 : the first version of the UMLS metathesaurus. In *Proceedings of the 13 th Annual Symposium on Computer Applications in Medical Care*, 1989. P. 483- 487

Annexe I

Liste des sigles utilisés :

AAT : Art and Architectural Thesaurus

CDD : Classification Décimale de Dewey

CDU : Classification Décimale Universelle

CLC : Chinese Library Classification

CMIT : Current Medical Information and Terminology

CT : Chinese Thesaurus

LCSH : Library of Congress Subject Headings

LIV : Legislative Indexing Vocabulary

MeSH : Medical Subject Headings

PDQ : the National Cancer Institute's Retrieval System

SHCL : Subject headings for children's Literature

SNOMED : Systematised Nomenclature of Medicine

TGM : LC Thesaurus of Graphic Materials