



68th IFLA Council and General Conference

August 18-24, 2002

Code Number: 007-122-F
Division Number: IV
Professional Group: *Classification et indexation*
Joint Meeting with: -
Meeting Number: 122
Simultaneous Interpretation: -

Recherche matière dans des ressources partagées : brève vue d'ensemble de récents développements

Martin Kunz

Die Deutsche Bibliothek
Frankfurt am Main
Allemagne

Résumé :

La recherche par sujets dans les ressources partagées est un problème actuel de la recherche en ligne d'informations bibliographiques. La construction de portails de ressources semblables n'est que la première étape, et la navigation qui suit dans les différentes interfaces de recherche n'est pas sans difficultés. Pour faciliter la recherche, il faut adapter ces différents ressources.

Les possibilités (normalisation / recherche croisée) et méthodes (automatique / intellectuelle) seront discutées. Ceci inclut une brève évaluation de l'avenir du travail terminologique multilingue :

- *approche classique (thésaurus multilingues)*
- *approche Internet (liens).*

Les développements récents dans les environnements mono- et multilingues seront présentés (MACS, CARMEN, Crosskonkordanz Wirtschaft).

1. La situation de départ

Jetons tout d'abord un rapide coup d'oeil en arrière, et rappelons-nous à quoi ressemblait le monde des bibliothèques il n'y a pas bien longtemps. L'espace germanophone en particulier était caractérisé, en ce qui concerne la recherche matière dans les bibliothèques, par une variété qui semblait inépuisable de procédés et de règles et ceci aussi bien dans le domaine de la recherche matière systématique par classification qu'analytique par langage alphabétique. Nous allons dans ce qui suit consacrer notre attention à cette dernière.

Nombre de ces méthodes étaient des solutions purement internes et chacune d'entre elles était naturellement meilleure que toutes les autres. *Tempi passati*. La grande majorité des bibliothèques de recherche et publiques de l'ensemble de l'espace germanophone utilise pour la recherche matière alphabétique un ensemble de règles unifié, à savoir les RSWK (*Regeln für den Schlagwortkatalog*). Avec cet ensemble de règles utilisé depuis maintenant plus de 15 ans, le SWD (*Schlagwortnormdatei*) s'est établi comme fondement de la recherche matière alphabétique. Il s'est développé pour l'indexation comme le thésaurus universel de l'espace germanophone, et s'est en outre étendu depuis longtemps au delà de son cercle de "clients" originel, les bibliothèques universitaires et les bibliothèques publiques. Des bibliothèques spécialisées travaillent aussi avec succès avec le SWD, même si ce n'est pas toujours en suivant la doctrine pure des RSWK.

La cible principale du vocabulaire de recherche SWD était naturellement les monographies, celles-ci constituant toujours la majeure partie des formes de publication. D'un autre côté, les différentes tâches de nos bibliothèques les amènent à indexer d'autres types de documents. Ainsi, et presque sans que cela ait été remarqué par le public des bibliothèques, des articles de périodiques destinés à des bibliographies régionales sont indexés avec le SWD dont les descripteurs ne sont plus seulement utilisés pour les publications imprimées, mais aussi pour celles d'autres médias, comme les objets de musée et les émissions de télévision.

A cette occasion il s'est révélé que les responsables du développement des RSWK et du SWD furent bien inspirés de s'en tenir, pour les questions de terminologie, c'est-à-dire ici pour les questions de la pré- ou de la post-coordination et de la représentation structurée du langage documentaire, de s'en tenir donc à ce qui a été établi dans les normes correspondantes par les spécialistes de la bibliothéconomie et de la documentation. Ainsi a pu naître un langage documentaire qui correspond dans sa structure et son contenu à l'état de l'art et qui grâce à cela est capable par la force des choses de trouver des intéressés aussi au delà du monde des bibliothèques. Le vocabulaire du SWD devrait par principe être utilisable pour tous les types de documents, il est au delà de l'étroitesse provinciale, et dépasse dans son universalité le profil d'une bibliothèque particulière.

2. Le problème de l'hétérogénéité

Ce qu'on appelle aujourd'hui le paysage mondial de l'information rend possible théoriquement de rechercher depuis n'importe où dans n'importe quelle banque de données bibliographiques, ces bases de données contenant d'ailleurs tous les types de documents. A leur tour, ces documents sont accessibles par sujet, selon des procédés d'indexation divers. Dans le domaine de l'indexation alphabétique, le SWD a une position dominante, mais pas de monopole. Il existe en outre, légitimement, des thésaurus spécialisés qui ont été développés par un travail de plusieurs années dans des institutions de documentation de rang national.

Même si l'utilisation d'un seul thésaurus universel pour toutes les institutions d'indexation et de fourniture de documents dans une aire linguistique donnée peut être un cas idéal, on ne peut pas s'attendre à ce que des établissements renommés réindexent maintenant leurs fonds. Nous devrions être heureux qu'au moins les bibliothèques se servent du même langage de documentation.

Jetons un coup d'oeil dans un tel thésaurus spécialisé (*BISp-Deskriptoren* du Bundesinstitut für Sportwissenschaft) et comparons-le au SWD. Nous trouvons, comme on pouvait s'y attendre, de nombreux accords, mais aussi quelques divergences. Au total, on peut cependant dire qu'il n'existe pas de concept qui ne se laisse pas rendre par l'un ou l'autre des vocabulaires d'indexation. Entre les formes retenues (vedettes) de chacun des langages documentaires, il

existe soit des identités (coïncidences), soit des équivalences 1:1 ou 1:X, ces dernières représentant des relations ET ou OU.

| | | |
|------------------------|--------------------------|--------------------|
| BISp-Deskriptoren | SWD | |
| <i>Abwehrtaktik</i> | <i>Abwehrtaktik</i> | <i>Identité</i> |
| <i>Alltagsobjekt</i> | <i>Alltagsgegenstand</i> | <i>Équivalence</i> |
| <i>Abweherschulung</i> | <i>Abwehr + Training</i> | <i>ET logique</i> |
| <i>Abwehr</i> | | |
| | <i>Abwehr</i> | <i>OU logique</i> |
| <i>Abwehrrspiel</i> | | |

Qu'il n'y ait pas d'équivalence dans une série de cas n'est pas non plus une surprise, le SWD est après tout un thésaurus encyclopédique et, d'un thésaurus spécialisé, on doit s'attendre à un vocabulaire plus spécifique, bien qu'il n'y ait pas d'obstacle à une reprise de ces descripteurs par le SWD dans certaines conditions qu'on peut caractériser comme une compatibilité terminologique. D'autre part, avec les projets que je voudrais montrer ci-après, il s'est révélé que c'est plutôt le SWD qui possède le vocabulaire professionnellement le plus spécifique et le plus vaste.

Mettre en relation de telles terminologies différentes était encore factuellement impossible il y a quelques années, mais, aujourd'hui, nous avons à notre disposition des techniques qui vous sont certainement suffisamment connues à partir d'Internet, mais qui aussi comportent des pièges et ont leurs limites. Que l'on nomme ce procédé "Crosskonkordanz" ou "Crosswalk" est en définitive indifférent. Pour simplifier, on peut dire qu'il s'agit de la création de liens entre les termes équivalents, représentant des concepts identiques, des deux thésaurus ; il s'agit de l'affiliation (un terme plus approprié ne me vient pas à l'esprit) de langages documentaires. Cela demande un système situé entre les thésaurus, et qui administre les liens existants, gère les modifications apportées à ceux-ci et l'introduction de nouveaux liens, et aide en même temps l'utilisateur lors de la navigation entre les ensembles de données respectifs.

3. Les projets

Dans ce qui suit, j'aimerais esquisser trois projets, qui sont semblables dans leur structure, deux d'entre eux sont monolingues (CARMEN, Crosskonkordanz Wirtschaft), le troisième est multilingue (MACS).

3.1 CARMEN

Une partie du Projet CARMEN consiste à relier le thésaurus IZT de l'Informationszentrum Sozialwissenschaften avec le SWD. La méthode par laquelle les équivalences sont déterminées et les liens établis est d'une séduisante simplicité : en partant de listes alphabétiques qui contiennent les vedettes-matière d'un domaine donné, les relations entre les deux thésaurus sont déterminées intellectuellement. Il n'y a pas très longtemps, une telle méthode aurait semblé complètement désuète, et chacun aurait considéré comme allant de soi de développer un nouveau moteur de recherche, un moteur de recherche meilleur que tous les précédents, la mère de tous les moteurs de recherche pour ainsi dire.

Mais l'euphorie sans limites en ce qui concerne les capacités se semblables produits semble disparue, disparue aussi la phase où en tout lieu on entendait (librement adapté de 1. Moses, 10) : "Allons, mettons-nous en route et bâtissons un moteur de recherche, afin de nous faire

un nom et de ne pas être dispersés". "La conception largement répandue selon laquelle le progrès rapide de la technique nous fournirait bientôt des moteurs de recherche qui, d'un seul coup, rendraient superflus le contrôle et la normalisation terminologiques tant détestés appartient au royaume de la magie ou au langage de la vente" (Christof Wolters: Ist die Schlagwortnormdatei für die Objektdokumentation im Museum geeignet? [Le SWD est-il approprié pour la documentation des objets de musée?] In: AKMB-News 1/1998).

Nous avons aussi rejeté dans le cadre de ce projet l'idée d'un ajustement automatique par un procédé encore à développer car, quoi qu'une telle méthode ait pu produire, chaque descripteur qui aurait été reconnu comme homographe aurait peut-être encore dû être vérifié intellectuellement pour savoir si le même contenu conceptuel correspondait bien aux mêmes lettres. Au lieu de cela, nous nous sommes mis vigoureusement au travail intellectuel et, au cours du premier semestre, avec l'aide d'une collaboratrice scientifique à mi-temps, nous avons traité 3 500 jeux de vedettes-matière de sorte que les relations d'équivalence entre le thésaurus de l'Informationszentrum Sozialwissenschaften, celui du Deutsche Institut für Pädagogische Forschung et le SWD soient établies et qu'elles puissent être entrées dans un système de gestion des liens.

Voici une page-exemple de concordance entre le thésaurus de l'Informationszentrum Sozialwissenschaften [IZT] et le SWD :

IZT-Thesaurus

SWD

| | | | |
|--------------------------------|----|---|--------------------------------|
| Rollenwandel | < | g | Rolle |
| Segregation | = | h | Segregation <Soziologie> |
| soziale Entwicklung | =+ | h | Gesellschaft + Entwicklung |
| soziale Integration | = | h | Soziale Integration |
| soziale Stabilität | < | g | Stabilität |
| sozialer Prozess | = | h | Sozialer Prozess |
| sozialer Wandel | = | h | Sozialer Wandel |
| soziokulturelle Entwicklung | = | h | Soziokultureller Wandel |
| sozioökonomische Entwicklung | = | h | Sozioökonomischer Wandel |
| Systemveränderung | =o | h | Systemveränderung |
| Systemveränderung | =o | h | Politischer Wandel |
| Transformation | <o | m | Politischer Wandel |
| Transformation | <o | m | Sozialer Wandel |
| Unterentwicklung | = | h | Unterentwicklung |
| Wertwandel | = | h | Wertwandel |
| wissenschaftlicher Fortschritt | = | h | Wissenschaftlicher Fortschritt |

Même si vous ne maîtrisez pas l'allemand, vous vous apercevez que les véritables concordances prédominent, ce qui est assez naturel puisque nous sommes ici dans un environnement unilingue. Les alternatives de descripteurs différents pour les mêmes concepts existent certes, cependant les spécialistes se décident le plus souvent pour les mêmes termes usuels. Au total, en comparant les thésaurus, il ressort ainsi environ 58% d'équivalences, dont 75% sont des identités. On a remarqué à cette occasion que le SWD dispose d'un vocabulaire spécialisé nettement plus étendu que les deux ensembles terminologiques professionnels.

3.2 Crosskonkordanz Wirtschaft

Comme pour le projet CARMEN, il s'agit ici d'établir des liens entre les descripteurs reconnus comme équivalents du Standard-Thesaurus Wirtschaft et du SWD.

Contrairement à CARMEN, nous exécuterons cette fois un contrôle automatique des doublons. Nous ne savons pas si cette méthode sera profitable, et nous ignorons par conséquent si elle nous économisera du travail intellectuel ou si elle peut conduire à une confusion plus grande. De toute manière, il est important pour nous d'acquérir à cette occasion de l'expérience pour des projets futurs, de façon à ne plus dépendre des théories subjectives.

3.3. MACS

L'année 1997 a marqué une coupure et une avancée qualitative dans l'existence du SWD, qui avait déjà 10 ans. Dans les dernières années, on n'a cessé de réclamer son ouverture vers les termes en langues étrangères. Dans les domaines pour lesquels le langage scientifique comme la langue de publication dans l'espace germanophone sont déterminés par l'anglais, des renvois en anglais ont été saisis depuis toujours dans le SWD, et de même, dans les domaines culturels, on a utilisé les termes italiens ou français usuels dans le langage spécialisé. L'acceptation de telles formes de renvois s'est faite au total en fonction des besoins d'un SWD germanophone. Mais les exigences de la littérature en langue étrangère et l'échange international des données de recherche par sujets dans un paysage de bibliothèques en réseau au niveau mondial n'avaient pas été pris en considération jusqu'ici.

Les premiers efforts du département de recherche matière de la DDB pour combler cette lacune d'informations remontent au début de la dernière décennie.

Grâce à une initiative récente de la bibliothèque nationale suisse, le projet MACS (Multilingual Access to Subjects) a pu alors être lancé. À côté de la bibliothèque nationale suisse participent à ce projet la Bibliothèque nationale de France, la British Library et la Deutsche Bibliothek. L'objectif de MACS est d'établir une relation entre les trois grandes listes d'autorité-matière LCSH, RAMEAU et SWD.

Le but immédiat du projet est d'établir, dans chaque liste et pour quelques domaines spécialisés choisis, les formes vedettes équivalentes à celles des deux autres listes d'autorité. Le travail s'est d'abord limité aux secteurs du sport et du théâtre, ainsi qu'à un choix de descripteurs très fréquemment utilisés.

Comme dans le projet CARMEN, les équivalences entre les descripteurs ont été établies au niveau intellectuel, sans l'aide de systèmes automatisés. Nous n'avons pas estimé qu'il nous revenait d'apporter au problème de la traduction automatique sa solution définitive en cherchant à développer un programme qui, pour chaque mot, reconnaisse sa correspondance terme à terme dans un autre langage. "Machine translation, does it exist? The only honest answer to this question is NO" (S. Krauwer: Machine Translation: State of the Art, Trends and User Perspective. In :TELRI-Proceedings of the first European Seminar "Language Resources for Language Technology". Budapest 1996).

| | | | | | |
|------------|---|---------------|---|------------------------------|--|
| sh85147274 | Women | frBN001838184 | Femmes | 040182029 | Frau |
| sh85147294 | Women -- Employment | frBN001521764 | Femmes -- Travail | 0401820292 + 04069349X | Frau + <i>Berufstätigkeit</i> |
| sh85147456 | Women authors | frBN001614243 | Femmes écrivains | 040533115 | Schriftstellerin |
| sh85147587 | Women in literature | frBN002484062 | Femmes -- Dans la littérature | 041136179 | Frau <Motiv> |
| sh85148133 | Work | frBN002059353 | Travail | 040025675 | Arbeit |
| sh85148146 | Work environment | frBN001764645 | Conditions de travail | 040026418 | Arbeitsbedingungen |
| sh85073639 | Working class | frBN001578199 | Classe ouvrière | 040687996 | Arbeiterklasse |
| sh85148201 | World history | frBN001555147 | Histoire universelle | 040791580 | Weltgeschichte |
| sh85148236 | World War, 1914- 1918 | frBN001617257 | Guerre mondiale (1914-1918) | 040791636 | Weltkrieg <1914- 1918> |
| sh85148273 | World War, 1939- 1945 | frBN002351158 | Guerre mondiale (1939-1945) | 04079167X | Weltkrieg <1939- 1945> |
| sh85148515 | World War, 1939- 1945 -- Underground movements | frBN001634192 | Guerre mondiale (1939-1945) frBN001634192 -- Mouvements de résistance | 04079167X + 040792625 | Weltkrieg <1939- 1945> + <i>Widerstand</i> |
| sh85149310 | Youth | frBN001555624 | Jeunesse | 040288595 | Jugend |

La création d'un thésaurus plurilingue complet n'est pas envisagée. Celui-ci requerrait la qualité de l'isomorphisme, c'est-à-dire que les équivalences devraient exister non seulement en ce qui concerne les vedettes, mais aussi en ce qui concerne la structuration et donc que pour chaque notion sur un même niveau hiérarchique il existe un équivalent correspondant dans chacun des langages associées. Pour cela, il est nécessaire d'adapter l'une à l'autre les structures des listes d'autorité nationales et, le cas échéant, de combler les lacunes existantes. Le produit fini complet pourrait remplacer en fin de compte les fichiers nationaux, puisqu'ils seraient un élément intégré de MACS.

Cette approche classique d'un thésaurus plurilingue n'a aucune chance d'être jamais réalisée. Après les expériences pratiques du projet MACS et un coup d'oeil sur les thésaurus spécialisés plurilingues existants, on peut se demander si les thésaurus plurilingues sous forme pure selon les prescriptions des normes nationales et internationales en vigueur ont encore un sens, et si tout simplement la dépense considérable nécessaire pour élaborer de tels moyens de documentation se justifie encore.

L'objectif à plus longue échéance de l'entreprise MACS est naturellement la saisie de tous les domaines. L'ouverture vers d'autres langues serait alors la prochaine étape logique.

La procédure visée pour MACS ne touchera pas les structures des listes d'autorités nationales respectives, pas plus qu'avec CARMEN l'IZT ne doit être modifié. Cela n'exclut pas qu'on en vienne au cours des travaux à des modifications sur les thésaurus respectifs, cependant il s'agit de corrections ou d'ajouts de précisions, qui, de toute manière, auraient été faites tôt ou tard, et apparaissent ici comme sous-produit d'une amélioration qualitative des différents fichiers.

L'utilisateur fait ses recherches dans sa langue maternelle ou dans un l'environnement linguistique qui lui est familier, et peut alors avec les équivalents qui sont établis, poursuivre son travail dans les bases de données bibliographiques qui y sont affiliées de cette manière.

La technique nécessaire ne doit pas être développée à nouveau, elle nous est familière avec Internet, les interfaces existent (Z39. 50) et tout ceci peut être utilisé pour la mise en relation des fichiers autorités avec les notices bibliographiques, pour autant que celles-ci soient accessibles sur le réseau.

4. Interface utilisateur

Quels que soient les types de documents enregistrés dans une base de données, qu'il s'agisse de publications conventionnelles ou électroniques, ou bien encore de films, d'objets de musée ou d'archives, qu'ils existent seulement avec une notice bibliographique "classique" ou avec Dublin Core, en version complète ou sous forme d'extraits, le problème de l'utilisateur demeure inchangé.

Comme auparavant, deux systèmes dans lesquels le savoir est organisé d'une manière différente se font face. Comme relation alphabétique entre ces systèmes, l'utilisateur d'une part et les bases de données bibliographiques d'autre part, un thésaurus, qui en plus de l'ordre dans le système "base de données" remplit deux fonctions :

1. Aide pour les recherches de l'utilisateur et
2. Navigation structurée dans les fichiers.

MACS relie grâce aux équivalences déterminées intellectuellement les contenus de bases de données bibliographiques qui utilisent pour leur indexation des vocabulaires contrôlés et qui montrent ainsi un fonds de données arrangé et structuré.

Il m'apparaîtrait par conséquent comme peu utile et même comme une régression d'installer - ce qui serait tout de même une alternative concevable à nos projets - sur des fonds de données de ce genre une procédure automatique avec des dictionnaires électroniques. Cela signifierait pour MACS que l'attribution conceptuelle univoque existante dans les systèmes respectifs serait dissoute, donc que l'utilisateur pourrait naviguer ainsi d'ambiguïté polysémique en ambiguïté polysémique.

Zug -> *anglais* : groove, pull, course, train, convoy drift, flight, tendency, draught, draft

Draft -> *allemand* : Tratte, Wechsel, Abteilung, Aufgebot, Aushebung, Nachschub, Entwurf, Skizze, Konzept, Zug

Wechsel -> *français* : échange, retraite, ...

Retraite -> *allemand* : Einsiedelei, Rückzug, Ruhestand, Wechsel, Zapfenstreich...

Si l'utilisateur connaît parfaitement sa terminologie spécialisée en langue étrangère, il n'a pas besoin d'un tel produit, et s'il ne la connaît pas bien, ou n'en est pas sûr, alors il ne l'aide pas. Un lien depuis la forme vedette d'un thésaurus vers son équivalent dans un autre pourrait toutefois probablement lui être utile.

Voici comme exemple simple la recherche pas à pas depuis le SWD par l'intermédiaire de MACS dans la bibliographique de la Bibliothèque nationale de France : supposons qu'un utilisateur recherche quelque chose sur le cyclisme en général. Incertain de la terminologie, l'utilisateur commence avec le mot de recherche "*Radfahren*", qui le conduit à la vedette Radsport. Il serait arrivé au même résultat s'il était entré dans le SWD avec le mot "*Fahrradsport*" ou seulement "*Fahrrad*" (les mots de recherche sont en italique) :

Fahrrad

Q : M

SYS 31.7

BF (EP) Velo

OB (TG) Zweirad

VB (TA) *Radfahren*

VB (TA) Velomobil

Radsport

Q : Sport-B

SYS 34.3

BF (EP) *Fahrradsport*

VB (TA) *Radfahren*

VB (TA) Radsportler

Après que notre utilisateur a examiné les titres du catalogue DDB, il trouve le lien vers MACS (je considère qu'on lui a communiqué avec des mots courts et à la portée de tous de quoi il s'agit là et ce qui l'attend). Il presse ainsi par curiosité sur les boutons disponibles et atterrit alors sur la page suivante, qui se présente naturellement pour l'utilisateur dans sa langue maternelle :

| | | |
|--|--|---|
| Suchwort <input type="text" value="radsport"/> Bemerkung: Es werden nur Schlagwörter gesucht | eingegeben Gewählte Sprache / Sprache auswählen <input type="text" value="Deutsch (SWD)"/> | Gewählte Bibliothek(en) / Bibliothek(en) auswählen <input type="text"/> <input checked="" type="checkbox"/> Swiss National Library <input checked="" type="checkbox"/> Bibliothèque nationale de France <input type="checkbox"/> Die Deutsche Bibliothek <input type="checkbox"/> The British Library Sie können eine oder mehrere Bibliotheken auswählen |
| <input type="text" value="Search"/> Suche in den ausgewählten Bibliotheken | | |

Puisque notre ami du cyclisme connaît l'importance de ce sport en France, l'offre de la BnF sur ce sujet l'intéresse en priorité :

Vos résultats de la recherche dans les bibliothèques...

Bibliothèque nationale de France : 31 réponses.

1. Il était une fois le cyclisme à Corbeil-Essonnes; un siècle d'histoire; Guy Caput, Roland Oberle;
2. Guide du cyclisme; manuel pratique et conseils de santé; docteur Gérard Porte;
3. Le livre d'or du cyclisme; 1995; Jean-Luc Gatellier; préf. de Laurent Jalabert;
4. L'année du cyclisme, 1996; Pierre Chany et Claude Droussent;
5. Le livre d'or du cyclisme; 1996; Jean-Luc Gatellier; préf. de Richard Virenque;
6. Cyclisme; 56 exercices et programmes; C. Carmichael, E. R. Burke; [trad. par Gwénaél Hubert];
7. Cyclisme; les conquérants de l'Arc-en-Ciel, 1927-1996; les champions du monde; par Claude Dassonville;
8. L'année du cyclisme 1997; Claude Droussent;
9. La fabuleuse histoire du cyclisme; Pierre Chany;
10. Le livre d'or du cyclisme 1997; Jean-François Quénet; préf. de Marc Madiot; photos, Graham Watson;

Ceci constitue certes un exemple simple, mais, puisque d'après notre expérience actuelle, dans plus de 60% des cas les équivalences sont en effet de 1:1, je ne considère pas du tout le processus comme éloigné de la réalité.

Une recherche structurée sur les formes vedettes du système familial est à coup sûr toujours plus simple à exécuter qu'un feuilletage dans un environnement de dictionnaire non structuré.

La fonction feuilletage dans MACS est clarifiée dans l'exemple suivant.

Un utilisateur voudrait voir toutes les vedettes dans lesquelles figure le mot (polysémique) "Training"

| | | |
|--|---|--|
| Type in the subject <input type="text" value="training"/> Note: Only subject headings are included in the search | In the chosen language <input type="text" value="English (LCSH)"/> | In the chosen Libraries <input type="checkbox"/> Swiss National Library <input checked="" type="checkbox"/> Bibliothèque nationale de France <input type="checkbox"/> Die Deutsche Bibliothek <input type="checkbox"/> The British Library Tip: You can select or deselect the libraries you want to search |
| <input type="text" value="Search"/> To search the selected libraries | | |

| | English (LCSH) | Deutsch (SWD) | Français (RAMEAU) |
|--------------------------|---------------------------------|------------------------------|------------------------------------|
| <input type="checkbox"/> | All terrain cycling -- Training | Training and Geländeradsport | Entraînement and Vélo tout terrain |
| <input type="checkbox"/> | Bicycle racing -- Training | Training and Radrennsport | Entraînement and Courses cyclistes |
| <input type="checkbox"/> | Cycling -- Training | Training and Radsport | Entraînement and Cyclisme |
| <input type="checkbox"/> | Employees -- Training of | Mitarbeiterschulung | Personnel -- Formation |

Par la commande feuilletage "Index", il obtient la liste (partiellement affichée ci-dessus) qui est un produit fourni automatiquement, tandis que les liens qu'elle contient ont été déterminés intellectuellement. La liste montre les problèmes auxquels est confronté un système de gestion inter-thésaurus, à savoir le traitement des équivalences 1:X qui résultent des usages terminologiques différents des langages de documentation impliqués.

Plus la terminologie est plutôt basée sur une base lexicographique, plus la relation est simple. Puisqu'il s'agit avec les LCSH de même que, bien que dans une moindre mesure, avec RAMEAU, de vocabulaires fortement pré-coordonnés, il faut fréquemment plusieurs descripteurs SWD pour construire des équivalences par les opérateurs logiques ET ou OU, qui constituent un moyen de résoudre, au moins de manière approchante, le problème des équivalences floues.

Lorsque le nombre des langues participantes augmente, la gestion des liens devient naturellement de plus en plus plus complexe - cela vaut aussi de manière semblable pour le lien des thésaurus spécialisés germanophones au SWD. Pour ne pas devenir ici des victimes

de l'explosion combinatoire, il faudrait envisager que les nouveaux langages documentaires ne se relient qu'à un langage de leur choix, de préférence à celui qui est le plus semblable en ce qui concerne la caractéristique de post- et/ou pré-coordination.

MACS se base sur le fait que l'utilisateur accède aux résultats d'indexation alphabétique intellectuelle par l'intermédiaire de thésaurus. Nous avons affaire à des stocks de données bibliographiques d'ampleur considérable. Des hiérarchies simples, telles qu'elles vous sont familières, par exemple dans Yahoo!, ne sont pas capables de maîtriser de telles masses, les thésaurus différenciant là fondamentalement plus fortement et les indexations complètes, telles que sont fournies par exemple par la méthode de l'indexation syntaxique, encore plus. Lorsque, sur les projets d'affiliation de fonds hétérogènes, le nombre des documents accessibles augmente encore, alors l'offre d'une interface utilisateur structurante, facilement possible car déjà présente de manière inhérente grâce à notre méthode d'indexation, devrait être une judicieuse nécessité.

5. Un bref bilan

Les trois projets montrent qu'il est possible d'établir des relations entre des fonds de données hétérogènes avec un effort intellectuel raisonnable et sans dépense technique trop grande. Nous avons vu comment des fonds qui sont indexés avec des langages de documentation différents, peuvent être reliés.

La condition pour une telle procédure devrait être qu'il s'agisse de fonds vastes et importants avec des langages documentaires complètement élaborés. Ce ne peut pas être le sens de notre méthode que de soutenir les terminologies privées de petites institutions, ou de contribuer à la construction de développements parallèles.