



68th IFLA Council and General Conference

August 18-24, 2002

Code Number: 007-122-G
Division Number: IV
Professional Group: Classification and Indexing
Joint Meeting with: -
Meeting Number: 122
Simultaneous Interpretation: -

Sachliche Suche in verteilten Ressourcen: ein kurzer Überblick über neuere Entwicklungen

(Subject retrieval in divided resources: a short overview about recent developments)

Martin Kunz

Die Deutsche Bibliothek
Frankfurt am Main,
Germany

Abstract:

Subject retrieval in divided resources is a current problem in the online research for bibliographic informations. Building portals for similar resources is only the first step, the following navigation in different search interfaces is not without difficulties. To make the retrieval easier it is necessary to accommodate these different resources.

These possibilities (standardisation versus Cross-walks) and the methods (automatic versus intellectual) will be discussed. This includes a short appraisal of the future of multilingual terminological work:

- The “classical” approach (Multilingual Thesauri),
- The “Internet” approach (linking)

The recent developments in mono- and multilingual environments will be presented (MACS, CARMEN, Crosswalk Economy).

Zusammenfassung:

Die sachliche Suche in verteilten Ressourcen ist ein gängiges Problem bei der Online-Suche nach bibliographischen Informationen. Der Aufbau von Portalen für vergleichbare Quellen ist nur ein erster Schritt, die anschließende Navigation in unterschiedlichen Such-Oberflächen jedoch nicht ohne Probleme. Zur Erleichterung des Retrievals ist eine Anpassung dieser unterschiedlichen Ressourcen erforderlich.

Die Möglichkeiten hierzu (Normierung versus Crosswalks) und die Methoden (automatisch versus intellektuell) werden diskutiert. Dies schließt eine kurze Abschätzung der zukünftigen multilingualen Terminologearbeit ein:

- Der „klassische“ Ansatz (Multilinguale Thesauri),
- Der „Internet“ Ansatz (Verlinkung)

Die neueren Entwicklungen in ein- und mehrsprachigen Umgebungen werden vorgestellt (MACS, CARMEN, Crosskonkordanz Wirtschaft).

1. Die Ausgangslage

Lassen Sie uns zunächst einen kurzen Blick zurückwerfen und erinnern wir uns daran wie die bibliothekarische Welt vor noch nicht allzu langer Zeit aussah. Gerade der deutschsprachige Raum war zumindest über viele Jahre hinsichtlich der bibliothekarischen sachlichen Erschließung gekennzeichnet durch eine schier unerschöpfliche kreative Vielfalt der unterschiedlichsten Verfahren und Regelwerke, und zwar sowohl im Bereich der klassifikatorischen als auch der verbalen Sacherschließung. Letzterem wollen wir im folgenden unsere Aufmerksamkeit zuwenden.

Zahlreiche dieser Methoden waren reine hausinterne Lösungen und jede davon war natürlich besser als alle anderen. *Tempi passati*. Die große Mehrheit der wissenschaftlichen und öffentlichen Bibliotheken des gesamten deutschsprachigen Raums nutzt für die verbale Sacherschließung ein einheitliches Regelwerk, nämlich die RSWK (Regeln für den Schlagwortkatalog). Mir diesem seit nunmehr über 15 Jahren praktizierten Regelwerk hat sich als Grundlage der verbalen Sacherschließung die SWD (Schlagwortnormdatei) etabliert. Die SWD hat sich als Indexierungsgrundlage zu **dem** Universalthesaurus des deutschen Sprachraums entwickelt und dabei ihren ursprünglichen Adressatenkreis, die Universitätsbibliotheken und die öffentlichen Bibliotheken, längst überschritten. Auch wissenschaftliche Spezialbibliotheken arbeiten erfolgreich mit der SWD, wenn auch nicht immer nach der reinen Lehre der RSWK.

Hauptziel des Erschließungsvokabulars der SWD waren naturgemäß die Monographien, diese machen nun einmal nach wie vor den überwiegenden Anteil aller Publikationsformen aus. Andererseits führen die unterschiedlichen Aufgaben unserer Bibliotheken zur Erschließung auch andersartiger Dokumententypen. So werden fast unbemerkt von der großen bibliothekarischen Öffentlichkeit schon seit längerem Zeitschriftenaufsätze für Regionalbibliographien mit dem Vokabular der SWD erschlossen und es werden die Deskriptoren der SWD nicht mehr nur zur Dokumentation gedruckter Veröffentlichungen sondern auch anderer Medienarten wie Museumsobjekte oder Fernsehsendungen eingesetzt. Dabei zeigt es sich, dass die für die Weiterentwicklung der RSWK und der SWD Zuständigen gut beraten waren, sich in Terminologiefragen, das heißt hier vor allem in den Fragen der Post- bzw. Präkoordination sowie in der strukturierten Darstellung der Dokumentationssprache an das zu halten, was in den einschlägigen Normen von Fachleuten aus dem Bibliotheks- und Dokumentationswesen festgelegt ist. Dadurch konnte eine Dokumentationssprache entstehen, die in Struktur und Inhalt dem state of the art entspricht, und die damit zwangsläufig auch jenseits der bibliothekarischen Welt ihre Interessenten zu finden vermag. Das Vokabular der SWD sollte grundsätzlich für alle Dokumententypen verwendbar sein, es ist jenseits provinzieller Enge und überschreitet in seiner Universalität das Profil einer einzelnen Bibliothek.

2. Das Problem der Heterogenität

Die heutige global genannte Informationslandschaft ermöglicht es theoretisch, von jedem Ort aus in jeder beliebigen bibliographischen Datenbank zu recherchieren, wobei in den Datenbanken alle Dokumententypen vertreten sind. Diese Dokumente sind wiederum mit jeweils unterschiedlichen

Erschließungsverfahren der sachlichen Suche zugänglich. Im Bereich der verbalen Inhaltserschließung hat die SWD zwar breitesten Zuspruch aber eben kein Monopol. Daneben existieren berechtigterweise Fachthesauri, die in jahrelanger Arbeit meist von Dokumentationseinrichtungen nationalen Ranges aufgebaut wurden. Auch wenn die Verwendung eines einzigen universellen Thesaurus für alle erschließenden und dokumentierenden Institutionen in einem gegebenen Sprachraum ein Idealfall sein mag, so kann nicht erwartet werden, dass renommierte Dokumentationseinrichtungen ihre Bestände nun neu indexieren. Wir sollten froh sein, dass zumindest die Bibliotheken sich derselben Dokumentationsprache bedienen.

Werfen wir einen kurzen Blick in einen solchen Fachthesaurus und vergleichen wir ihn mit der SWD. Erwartungsgemäß finden wir zahlreiche Übereinstimmungen aber auch einige Abweichungen. Insgesamt lässt sich aber sagen, dass es praktisch keinen Begriff gibt, der sich nicht mit dem einen oder anderen Indexierungsvokabular wiedergeben ließe. Zwischen den Vorzugsbenennungen der jeweiligen Dokumentationsprachen bestehen also entweder Kongruenzen, 1:1-Äquivalenzen oder 1:X-Äquivalenzen, wobei letztere UND- bzw. ODER-Beziehungen verkörpern.

BISp-Deskriptoren (Bundesinstitut für Sportwissenschaft)	SWD	
Abwehrtaktik	Abwehrtaktik	Kongruenz
Alltagsobjekt	Alltagsgegenstand	Äquivalenz
Abweherschulung	Abwehr + Training	logisches UND
Abwehr	Abwehr	logisches ODER
Abwehrspiel		

Dass es in einer Reihe von Fällen keine Äquivalenz gibt ist auch keine Überraschung, die SWD ist schließlich ein Universalthesaurus und von einem Fachthesaurus sollte man ein spezifischeres Vokabular erwarten, wengleich einer Übernahme dieser Deskriptoren durch die SWD unter bestimmten Voraussetzungen, die man als terminologische Kompatibilität charakterisieren mag, nichts im Wege steht. Andererseits hat sich bei den Projekten, die ich im folgenden zeigen möchte, erwiesen, dass eher die SWD das fachlich spezifischere, umfangreichere Vokabular besitzt. Solche unterschiedlichen Terminologiebestände aufeinander zu beziehen war noch vor wenigen Jahren faktisch unmöglich, heute stehen uns Techniken zur Verfügung, die Ihnen aus dem Internet hinreichend bekannt sein dürften, die aber auch ihre Tücken und Grenzen haben. Ob man dieses angestrebte Verfahren nun als Crosskonkordanz oder als Crosswalk bezeichnet ist letztlich egal, vereinfachend gesagt geht es um die Herstellung von Links zwischen den äquivalenten, gleiche Begriffe repräsentierenden Benennungen der beiden Thesauri, es geht um die Affiliation (ein besseres Wort fällt mir dafür nicht ein) von Dokumentationsprachen. Dazu bedarf es eines zwischen den Thesauri liegenden Systems, welches die vorhandenen Links verwaltet, Änderungen an diesen sowie die Einführung neuer Links steuert und zugleich die Benutzer bei der Navigation zwischen den jeweiligen Datenbeständen unterstützt.

3. Die Projekte

Im folgenden möchte ich drei Projekte skizzieren, die in ihrer Struktur ähnlich sind, zwei davon sind monolingual (CARMEN, Crosskonkordanz Wirtschaft), das dritte ist multilingual (MACS).

3.1. CARMEN

Ein Teilbereich des CARMEN-Projekts befasst sich mit der Verknüpfung des Thesaurus des Informationszentrums Sozialwissenschaften (IZT) mit der SWD. Die Methode, mittels derer die Äquivalenzen bestimmt und die Links erstellt werden, ist von bemerkender Schlichtheit: ausgehend von alphabetischen Listen, die das Schlagwortmaterial aus einem spezifischen Sachgebiet enthalten, werden die Beziehungen zwischen beiden Thesauri intellektuell bestimmt. Vor nicht allzu langer Zeit

wäre eine solche Vorgehensweise wohl antiquiert erschienen und jeder hätte es für selbstverständlich gehalten, eine neue Suchmaschine zu entwickeln, eine Suchmaschine, besser als alle bisherigen, die Mutter aller Suchmaschinen gewissermaßen. Aber die grenzenlose Euphorie über die Leistungsfähigkeit solcher Produkte scheint dahin zu sein, dahin auch die Phase, als es frei nach 1.Moses,10 allenthalben hieß „wohlan lasset uns hingehen und eine Suchmaschine bauen, dass wir uns einen Namen machen und nicht zerstreuet werden“. „Die weit verbreitete Auffassung, dass der rasante Fortschritt der Technik uns schon in nächster Zukunft „Suchmaschinen“ bescheren würde, die mit einem Schlag die ungeliebte terminologische Kontrolle und Standardisierung unnötig machen würden, gehört in das Reich der Magie oder in die Sprache des Verkaufs“ (Christof Wolters: Ist die Schlagwortnormdatei für die Objektdokumentation im Museum geeignet? In: AKMB-News 1/1998). Auch die Idee eines automatischen Abgleichs mit Hilfe eines ja erst noch zu entwickelnden Verfahrens haben wir im Rahmen dieses Projektes verworfen, denn was auch immer eine solche Methode hätte leisten können, jeder Deskriptor, der z.B. maschinell als zeichengleich identifiziert worden wäre, hätte möglicherweise noch einmal intellektuell darauf hin überprüft werden müssen, ob den nämlichen Buchstaben auch der gleiche begriffliche Inhalt eignet. Statt dessen haben wir uns also rüstig und intellektuell an die Arbeit gemacht und im ersten Halbjahr mit einer halben wissenschaftlichen Mitarbeiterin 3.500 Schlagwortsätze so bearbeitet, dass die äquivalenten Relationen zwischen dem Thesaurus des Informationszentrums Sozialwissenschaften, dem Thesaurus des Deutschen Instituts für Pädagogische Forschung und der SWD bestimmt sind und sie in ein Linkverwaltungssystem eingespeist werden können.

Hierzu eine Beispielseite aus der Konkordanz zwischen dem Thesaurus des Informationszentrums Sozialwissenschaften und der 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

--	--	--	--

Auch wenn Sie des Deutschen nicht mächtig sind, so erkennen Sie doch, dass echte Kongruenzen überwiegen. Da wir hier in einer monolingualen Umgebung sind, liegt dies in der Natur der Sache. Die Alternativen zu unterschiedlichen Deskriptoren für die gleichen Begriffe sind zwar gegeben, dennoch entscheiden sich die Fachleute meist für die selben gebräuchlichen Benennungen. Insgesamt ergaben sich bei einem Vergleich der Thesauri ca. 85 % Äquivalenzen, davon ca. 75 % kongruent. Dabei fiel auf, dass die SWD über ein deutlich umfangreicheres Fachvokabular verfügt als die beiden fachlichen Terminologiebestände.

3.2. Crosskonkordanz Wirtschaft

Wie beim Projekt Carmen geht es auch hier um die Erstellung von Links zwischen den als äquivalent erkannten Deskriptoren des Standard-Thesaurus Wirtschaft und der SWD. Anders allerdings als bei Carmen werden wir diesmal einen ersten Dublettencheck automatisch durchführen. Wir wissen nicht, ob dieses Verfahren sinnvoll sein wird, ob es uns also intellektuelle Arbeit einspart oder ob es möglicherweise zu größerer Konfusion führt. Auf jeden Fall ist es uns wichtig, hier für spätere Projekte Erfahrungen zu gewinnen, damit wir dann nicht mehr auf subjektive Theorien angewiesen sind.

3.3. MACS

Das Jahr 1997 markierte einen Einschnitt und einen qualitativen Fortschritt in dem nunmehr über zehnjährigen Bestehen der SWD. In den letzten Jahren wurde immer wieder eine Öffnung der SWD für fremdsprachige Bezeichnungen gefordert. In den Bereichen, in denen die Wissenschaftssprache und auch die Publikationssprache im deutschen Sprachraum überwiegend vom Englischen bestimmt ist, wurden schon immer englischsprachige Verweisungen in der SWD erfasst, ebenso wurde in den Kulturwissenschaften mit in der Fachsprache gebräuchlichen italienischen oder französischen Bezeichnungen verfahren. Die Aufnahme solcher Verweisungsformen orientierte sich dabei insgesamt an den Bedürfnissen einer deutschsprachigen SWD. Bislang nicht berücksichtigt waren die Anforderungen fremdsprachiger Literatur und der internationale Austausch von Sacherschließungsdaten in einer global vernetzten Bibliothekslandschaft. Erste Bemühungen der Abteilung Sacherschließung der DDB, diese Informationslücke zu schließen, reichen bis zum Anfang des letzten Jahrzehnts zurück. Dank einer neuerlichen Initiative der Schweizerischen Landesbibliothek konnte dann das Projekt MACS (**M**ultilingual **A**ccess to **S**ubject **H**eadings) gestartet werden. An diesem Projekt nehmen neben der Schweizerischen Landesbibliothek die Bibliothèque Nationale de France, die British Library und die Deutsche Bibliothek teil. Ziel von MACS ist es, eine Verbindung zwischen den drei umfangreichen Schlagwortnormdateien LCSH, RAMEAU und SWD zu erarbeiten.

Nahziel des Projekts ist es, in jeder Normdatei in einigen ausgewählten Fachgebieten äquivalente Vorzugsbezeichnungen der anderen Normdateien entsprechend gekennzeichnet nachzuweisen. Zunächst beschränkte sich die Arbeit auf die Bereiche Sport und Theater, sowie auf eine Auswahl besonders häufig benutzter Deskriptoren. Die Äquivalenzen zwischen den Deskriptoren wurden wie auch im CARMEN-Projekt intellektuell erarbeitet, ohne Unterstützung automatisierter Systeme. Wir haben unsere Aufgabe nicht darin gesehen, das Problem des automatischen Übersetzens seiner endgültigen Lösung zu zuführen, indem wir ein Programm zu entwickeln trachteten, welches an einem einzelnen Wort dessen eineindeutige Entsprechung in einer anderen Sprache erkennt. „Machine translation, does it exist? The only honest answer to this question is NO“ (S. Krauer: 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 + Berufstätigkeit
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) -- Mouvements de résistance	04079167X + 040792625	Weltkrieg <1939-1945> + Widerstand
sh85149310	Youth	frBN001555624	Jeunesse	040288595	Jugend

Die Herstellung eines kompletten multilingualen Thesaurus ist nicht beabsichtigt. Dieser bedürfte der Eigenschaft der Isomorphie, d.h., dass Äquivalenzen nicht nur hinsichtlich der Vorzugsbenennungen, sondern auch hinsichtlich der Strukturierung bestehen müssen, dass also zu jedem Begriff auf gleicher Hierarchiestufe ein entsprechendes Äquivalent in jeder der beteiligten Sprachen vorhanden sein muss. Hierzu ist es erforderlich, die Strukturen der nationalen Normdateien einander anzupassen und gegebenenfalls vorhandene Lücken zu schließen. Das vollständige Endprodukt würde letztlich die nationalen Dateien ersetzen können, da sie integrierter Bestandteil von MACS wären. Dieser klassische Ansatz eines multilingualen Thesaurus hat keinerlei Chancen, jemals realisiert zu werden. Nach den praktischen Erfahrungen aus dem MACS Projekt und einem Blick in existierende multilinguale Fachthesauri ist doch große Skepsis angebracht, ob multilinguale Thesauri in reiner Form nach den Vorgaben der einschlägigen nationalen und internationalen Normen noch sinnvoll sind, ob sich also der erhebliche Aufwand überhaupt noch lohnt, solche Dokumentationsmittel zu erarbeiten.

Fernziel des Unternehmens MACS ist natürlich die Erfassung aller Sachgebiete. Die Öffnung für weitere Sprachen wäre dann der nächste logische Schritt.

Das für MACS angestrebte Verfahren wird die Strukturen der jeweiligen nationalen Normdateien nicht tangieren, genauso wenig wie bei CARMEN der IZT modifiziert werden muss. Das schließt nicht aus, dass es im Verlauf der Arbeiten zu Veränderungen an den jeweiligen Thesauri kommt, dabei handelt es sich aber um Korrekturen oder Präzisierungen, die ohnehin früher oder später erfolgt wären und hier quasi als Abfallprodukt eine qualitative Verbesserung der einzelnen Dateien nach sich ziehen.

Der Benutzer recherchiert in seiner Muttersprache oder in der ihm vertrauten sprachlichen Umgebung und kann dann mit den dort nachgewiesenen Äquivalenten in den solcherart affilierten bibliographischen Datenbanken weiterarbeiten. Die nötige Technik muss nicht neu entwickelt werden, sie ist uns aus dem Internet vertraut, die Schnittstellen sind vorhanden (Z39.50) und all dies kann für die Verbindung zwischen den Normdateien und den Titeldaten - vorausgesetzt, diese sind im Netz zugänglich - eingesetzt werden.

4. Benutzeroberfläche

Gleichgültig, welche Dokumentarten in einer Datenbank verzeichnet sind, ob konventionelle oder elektronische Publikationen oder ob es sich um Filme, museale Objekte oder Archivalien handelt, ob sie nun nur mit einer „klassischen“ Titelaufnahme oder mit Dublin Core, ob in Volltextversion oder in

Auszügen vorliegen, das Problem des Benutzers ist immer noch unverändert. Nach wie vor stehen sich zwei Systeme, in denen auf unterschiedliche Weise Wissen organisiert ist, gegenüber. Als verbale Verbindung zwischen diesen Systemen, dem Benutzer auf der einen und bibliographischen Datenbanken auf der anderen Seite, dient ein Thesaurus, der neben der Ordnung innerhalb des Systems Datenbank zweierlei Funktionen erfüllt:

1. Unterstützung bei der Anfrage des Benutzers und
2. Strukturierte Navigation in den Dateien.

MACS verknüpft über intellektuell bestimmte Äquivalenzen die Inhalte von bibliographischen Datenbanken, die für ihre Inhaltserschließung terminologisch kontrollierte Vokabularien einsetzen und die also einen geordneten, strukturierten Datenbestand vorweisen. Es erscheint mir deshalb als wenig hilfreich bzw. als ein Rückschritt, wenn über derartige Datenbestände ein automatisches Verfahren – das wäre zu unseren Vorhaben ja immerhin eine denkbare Alternative - mit elektronischen Wörterbüchern gelegt wird. Für MACS würde das bedeuten, dass die in den jeweiligen Systemen vorhandene eindeutige begrifflichen Zuordnung aufgelöst würde, der Benutzer also von Mehrdeutigkeit zu Mehrdeutigkeit navigieren dürfte.

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

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

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

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

Sollte der Benutzer seine fremdsprachige Fachterminologie perfekt kennen, dann braucht er ein solches Produkt nicht, und wenn er sie nicht gut kennt bzw. unsicher ist, dann hilft es ihm nicht. Wohl aber könnte ihm ein Link von der Vorzugsbenennung des einen Thesaurus zu deren Äquivalent in einem anderen dienlich sein. Hierzu als ausgesucht einfaches Beispiel die schrittweise Suche von der SWD über MACS in den Titelbestand der Bibliothèque Nationale de France: Nehmen wir an, ein Benutzer sucht etwas über den Radsport im allgemeinen, unsicher ob der Terminologie beginnt er mit dem Suchwort Radfahren, dieses führt ihn auf den Deskriptor Radsport, ähnlich wäre er dorthin gekommen, wenn er mit dem Wort Fahrradsport oder nur Fahrrad in die SWD eingestiegen wäre (die Suchwörter sind hervorgehoben):

Fahrrad

Q: M

SYS 31.7

BF Velo

OB Zweirad

VB **Radfahren**

VB Velomobil

Radsport

Q: Sport-B

SYS 34.3

BF **Fahrradsport**

VB **Radfahren**

VB Radsportler

Nachdem unser Benutzer die Titel der DDB in Augenschein genommen hat, findet er den Link zu MACS, wobei ich davon ausgehe, dass ihm mit kurzen und allgemeinverständlichen Worten mitgeteilt wurde, worum es sich da handelt und was ihn erwartet. Er drückt also neugierig auf den Button und

landet auf folgender Seite, die sich dann für den jeweiligen Nutzer natürlich in dessen Muttersprache präsentiert:

<p>Suchwort eingeben</p> <input type="text" value="radsport"/> <p><i>Bemerkung: Es werden nur Schlagwörter gesucht</i></p>	<p>Gewählte Sprache / Sprache auswählen</p> <input type="text" value="Deutsch (SWD)"/>	<p>Gewählte Bibliothek(en) / Bibliothek(en) auswählen</p> <input type="text"/> Swiss National Library <input checked="" type="checkbox"/> Bibliothèque nationale de France <input type="text"/> Die Deutsche Bibliothek <input type="text"/> The British Library <p><i>Sie können eine oder mehrere Bibliotheken auswählen</i></p>
<input type="button" value="Search"/>	<p>Suche in den ausgewählten Bibliotheken</p>	
<input type="button" value="Browse"/>	<p>Zeige die verknüpften Übersetzungen an</p>	

Da unser Radsportfreund um die Bedeutung dieser Sportart in Frankreich weiß, interessiert ihn das diesbezügliche Angebot der BNdF vorrangig:

Ihre Suchergebnisse in den Bibliotheken...
 Bibliothèque nationale de France: 31 hits.

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; \[trad. par GwénaÉel Hubert\];](#)
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;](#)

Dies ist zugegeben ein einfaches Exempel, aber da nach unserer bisherigen Erfahrung in über 60 % aller Fälle die Äquivalenzen tatsächlich 1:1 ausfallen, halte ich den Vorgang für durchaus nicht idealiter realitätsfern.

Eine strukturierte Suche über die Vorzugsbenennungen des vertrauten Systems ist jedenfalls allemal einfacher zu bewerkstelligen als ein Browsing in einer unstrukturierten Wörterbuch-Umgebung. In folgendem Beispiel sei die Browsing-Funktion in MACS verdeutlicht. Ein Benutzer möchte alle Schlagwörter sehen, in denen das (polyseme) Wort „Training“ vorkommt:

<p>Type in the subject</p> <input type="text" value="training"/>	<p>In the chosen language</p> <input type="text" value="Français (RAMEAU)"/>	<p>In the chosen Libraries</p> <input type="text"/> Swiss National Library
---	---	---

<i>Note: Only subject headings are included in the search</i>		<input type="checkbox"/> Bibliothèque nationale de France <input type="checkbox"/> Die Deutsche Bibliothek <input type="checkbox"/> The British Library <i>Tip: You can select or deselect the libraries you want to search</i>
<input type="button" value="Search"/>	To search the selected libraries	
<input type="button" value="Browse"/>	To view possible translations	

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

Über das Kommando „Browse“ erhält er dann die obige auzugsweise wiedergegebene Liste, sie ist ein automatisch erstelltes Produkt, während die in ihr enthaltenen Links intellektuell bestimmt wurden. Die Liste zeigt die Probleme, denen sich ein Interthesaurus Management System stellen muss, nämlich die Behandlung von 1:x-Äquivalenzen, die sich aus den unterschiedlichen terminologischen Gepflogenheiten der involvierten Dokumentationssprachen ergeben. Je eher die Terminologie auf lexikographischer Grundlage beruht, desto einfacher ist die Verbindung. Da es sich sowohl bei den LCSH als auch wenngleich in geringerem Ausmaß bei RAMEAU um sehr präkoordinierte Vokabularien handelt, bedarf es häufig mehrerer SWD-Deskriptoren, um durch UND- oder ODER-Logiken Äquivalenzen herzustellen, letztere sind auch ein Mittel, dem Problem unscharfer Äquivalenzen zumindest näherungsweise zu Leibe zu rücken.

Wenn die Anzahl der teilnehmenden Sprachen zunimmt, dann wird die Verwaltung der Links natürlich zunehmend komplexer – dies gilt in ähnlicher Weise auch für die Anbindung deutschsprachiger Fachthesauri an die SWD. Um hier nicht Opfer der kombinatorischen Explosion zu werden, wäre zu erwägen, dass neue Dokumentationssprachen sich nur an eine Sprache ihrer Wahl andocken, vorzugsweise an die Sprache, die hinsichtlich des Merkmals Post- bzw. Präkoordination der Eigenen am ähnlichsten ist.

MACS basiert darauf, dass der Benutzer sich die Ergebnisse intellektueller verbaler Inhaltserschließung über Thesauri erschließt. Wir haben es mit bibliographischen Datenbeständen von erheblichem Umfang zu tun. Schlichte Hierarchien, wie sie Ihnen etwa aus Yahoo geläufig sind, werden dieser Mengen nicht Herr, Thesauri differenzieren da wesentlich stärker und komplette

Indexate, wie sie etwa nach der Methode des syntaktischen Indexierens erstellt werden, noch mehr. Wenn nun über die Projekte zur Affiliation heterogener Bestände die Zahl der zugänglichen Dokumente noch einmal zunimmt, dann sollte das ohne weiteres möglich, da über unsere Indexierungsmethode bereits inhärent vorhandene Angebot einer strukturierenden Benutzeroberfläche eine einsichtige Notwendigkeit sein.

5. Ein kurzes Fazit

Die drei Projekte zeigen, dass es möglich ist, mit überschaubarem intellektuellen Aufwand und ohne allzu großen technischen Aufwand Verbindungen zwischen heterogenen Datenbeständen herzustellen. Wir haben gesehen, wie Bestände, die mit unterschiedlichen Dokumentations Sprachen erschlossen sind, miteinander verbunden werden können. Voraussetzung für ein solches Verfahren sollte sein, dass es sich um umfassende und bedeutende Bestände mit vollständig elaborierten Dokumentations Sprachen handelt. Es kann nicht der Sinn unserer Methode sein, private Terminologien kleiner Institutionen zu unterstützen oder zum Aufbau von Parallelentwicklungen beizutragen.