

Einführung in die verbale Inhaltserschließung für Bibliotheken

Dr. Gerhard Stumpf

Kurzfassung des Skripts zum Kompaktkurs Juni 2002

Einführung in die verbale Sacherschließung

Während sich die **klassifikatorische Erschließung künstlicher Indexierungssysteme** mit hierarchisch strukturierten Ordnungsprinzipien bedient, beruht die **verbale Inhaltserschließung** auf **natürlich-sprachlichen** (ohne methodische Normierung) **oder natürlich-sprachlich basierten** (mit Terminologiekontrolle) **Indexierungsmethoden**.

Die **verbale Inhaltserschließung (im Folgenden, dem nach wie vor gängigen bibliothekarischen Sprachgebrauch gemäß: Sacherschließung)** umfasst insbesondere die "klassische" **Schlagwortvergabe, die Stichwortvergabe und die Thesauruserstellung**.

Die **verbale Sacherschließung in Bibliotheken** unterscheidet sich von der im **Dokumentationswesen** herkömmlich vor allem dadurch, dass Bibliotheken im wesentlichen die monographische Literatur erschließen, Dokumentare sich aber auch und vor allem der unselbständig publizierten Literatur widmen. Im Dokumentationswesen wird im allgemeinen eine viel größere Erschließungstiefe erreicht.

In der Regel sind die Erschließungsverfahren für **universal ausgerichtete Bibliotheken** anders als für **Spezialbibliotheken**. Die in diesem Kurs vorgestellten Verfahren sind vorwiegend universal ausgerichtet und fächerübergreifend.

In den großen Verbunddatenbanken und Metakatalogen ist allerdings längst eine signifikante Verschiebung eingetreten. Die bibliothekarisch erschlossenen Materialien beschränken sich nicht mehr auf selbstständig erschienene Dokumente, sondern umfassen zunehmend auch unselbstständig erschienene Aufsätze (vor allem aus der regionalbibliographischen und der spezialbibliothekarischen Erschließung). Außerdem werden im Bereich der elektronischen Ressourcen zunehmend auch „ongoing publications“ und dynamische Websites inhaltlich erschlossen.

Die entstehenden **heterogenen Datenwelten** erfordern eine neue Aufgabendefinition der bibliothekarischen Inhaltserschließung, bieten aber auch die Chance, diese Leistung der Bibliotheken zu einem unverzichtbaren qualitativen Fundament für Datenbanken und Metadaten bibliographischer Art zu machen. Nach dem sog. **Schalenmodell**¹ gruppieren sich um einen Kern von Daten mit hochwertiger Erschließung verschiedene Schichten mit Sacherschließung von abgestufter Qualität bis hin zu Daten ohne jede Erschließung. Das Retrieval in einem solchen heterogenen Datenpool muss auf alle relevanten Dokumente zielen, wobei die bei einem Teil der Dokumente vorhandene gute Erschließung z.B. durch statistische und probabilistische Verfahren das Auffinden der weniger gut erschlossenen Dokumente erleichtert. Die in einem solchen Pool zusammengeführten Schichten könnten etwa folgende Charakteristika haben:

	Hochwertige Erschließung	▲▲▲▲▲ Dokumente mit syntaktischen Indexaten (z.B. RSWK)
		▲▲▲▲▲ Dokumente mit kontrollierten Deskriptoren (z.B. SWD)
		▲▲▲▲▲ Dokumente mit frei vergebenen Deskriptoren
		▲▲▲ Maschinell indexierte Titelaufnahmen
		▲▲ Mit Paratext (z.B. Abstract, Inhaltsverzeichnisse, Klappentexte) angereicherte Titelaufnahmen (mehr suchbare Stichwörter)
		▲ Titelaufnahmen ohne Inhaltserschließung (einzige verwendbare Elemente: Titelstichwörter)
	Minderwertige Erschließung	

¹ Vgl. etwa <http://www.mathematik.uni-osnabrueck.de/projects/slot3/workshop98II/KS.html>

Absolute **Konsistenz** (durchgängig gleich verwendetes Vokabular bei gleichmäßiger Erschließungstiefe) wird in großen Verbund- und Metadatenbanken immer weniger erreichbar sein. Es ist jedoch für Benutzer von großem Vorteil, wenn durch Hinterlegung von Normdateien und Harmonisierung der zu Grund liegenden Regeln eine möglichst weit gehende Konsistenz erzeugt wird. Außerdem kann es ein Vorteil für die Lösung vieler Suchprobleme sein, wenn die Suche in einem konsistent erschlossenen Teilsegment möglich ist, z. B. für einen ersten Überblick.

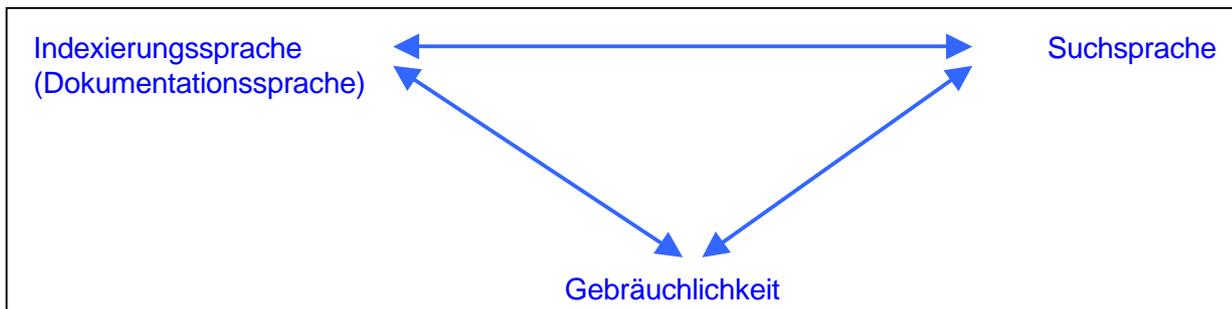
Die einzelne logische Einheit oder die Gesamtheit der Indexierungsbegriffe, die einem Dokument bzw. einer formalen Dokumentbeschreibung (= Katalogisat) zugewiesen werden, kann man als **Indexat** bezeichnen. Indexate sind Ergebnisse eines Indexierungsvorgangs; dieser kann intellektueller oder maschineller Art sein.

Aktive Inhaltserschließung (Indexierung) hat ihre größte Wirkung in Datenräumen aus heterogenem Material, wenn das Vokabular in den Dokumenten selbst nicht normiert ist.

Unter „**Schlagwortkatalog**“ kann man nicht nur den konventionellen Katalog verstehen, sondern auch die **Gesamtheit der verbalen Zugriffsmöglichkeiten auf den Dokumentinhalt in einem OPAC**, einschließlich der Titelaufnahmen.

Im Unterschied zum Systematischen Katalog (SYK), dessen Fachordnung häufig auch für die Aufstellung benutzt wird, wird der Schlagwortkatalog (SWK) **in jedem Fall standortunabhängig** geführt. Das bedeutet, dass Änderungen bei der Aufstellung der Bestände und Änderungen an einem SWK sich gegenseitig nicht tangieren.

Die **wesentliche Schwierigkeit** der verbalen Sacherschließung liegt in der verbalen Erfassung des Sachinhalts des zu erschließenden Werkes, d.h. in der **Schlagwortbildung**. Die Aufgabe besteht darin, vom Sachinhalt des Dokuments ausgehend, einen Ausdruck zu finden, unter dem der Benutzer das Dokument voraussichtlich suchen dürfte.



Von den **intellektuell vergebenen Schlagwörtern** klar zu unterscheiden sind die **Stichwörter** aus dem Hauptsachtitel oder anderen Titeln **sowie maschinelle Indexierung**, die Wörter aus Titeln oder weiteren Elementen des Dokuments wie Abstract, Inhaltsverzeichnis, Register u.a. verwendet.

Terminologie

Im Bereich der Sacherschließung ist eine **divergierende Terminologie** verbreitet. Dokumentations- und Bibliothekswesen haben sich in ihrer Fachsprache auseinander entwickelt.

Zur **Vereinheitlichung der Terminologie** trägt heute u.a. die **DIN 2330** (erste Ausgabe 1979, neueste Ausgabe 1993) bei, die auch als deutscher Beitrag in die entsprechende internationale Norm ISO 704 eingebracht wurde.

Ein Mindestmaß an **methodischer Übereinstimmung** und **Indexierungsqualität** bei den Indexierungsverfahren soll durch die **DIN 31623** (erste Ausgabe 1983, neueste Ausgabe 1988) erreicht werden.

So wurden auch für die **Begriffsbestimmungen in den RSWK die einschlägigen DIN-Normen** (DIN 1463 Thesauri, DIN 2330 Begriffe und Benennungen, DIN 31623 Indexierung und inhaltliche Erschließung von Dokumenten, DIN 31631 Kategorienkatalog für Dokumente) herangezogen.

Grundlegende Definitionen gemäß den RSWK

Die **Schlagwortkatalogisierung** basiert **auf der natürlichen Sprache**. Sie folgt festgelegten Regeln für die Ansetzung der Schlagwörter (vgl. § 9) sowie für deren Verwendung zur Wiedergabe eines Dokumenteninhalts. Die Schlagwörter werden in der überregionalen Schlagwortnormdatei (SWD) bzw. bei lokaler Anwendung in einer Schlagwortdatei geführt.

Die **Schlagwortnormdatei (Schlagwortdatei)** besteht aus **Schlagwortsätzen**. Diese enthalten **Felder für Ansetzungsformen und Verweisungen sowie ergänzend Codierungen** (einschließlich der Notation einer nationalen oder lokalen Klassifikation), **Bemerkungen und Verwaltungsinformationen** (vgl. § 7,4).

Bei der Zuordnung der Schlagwörter zu den Dokumenten sind **grundsätzlich zwei strukturell unterschiedliche Verfahren möglich**:

1. Die Verwendung einzelner nicht verknüpfter Schlagwörter, die unabhängig von documentspezifischen Beziehungen gleichrangig nebeneinander gestellt werden (**gleichordnende Indexierung**)
2. Die Verknüpfung einzelner Schlagwörter zu **Schlagwortketten**, die den dokumentenspezifischen Zusammenhängen entsprechen (**syntaktische Indexierung**)

Im Unterschied zum gleichordnenden Indexieren, bei dem die Schlagwörter gleichwertig nebeneinander stehen, werden bei der syntaktischen Indexierung nach RSWK die Schlagwörter, die einen Gegenstand eines Dokumentes repräsentieren sollen, in einer bestimmten Reihenfolge (vgl. §15) angeordnet. Behandelt ein Dokument mehrere Gegenstände, sind entsprechend mehrere **Schlagwortketten** zu bilden (DIN 31623 3.Teil 7.2).

Dokument: Als Dokument wird die vorliegende im Schlagwortkatalog nachzuweisende Ausgabe eines Werkes bezeichnet. Je nach Erscheinungsform sind bibliographisch selbständige Dokumente und bibliographisch unselbständige Dokumente zu unterscheiden (vgl. § 3).

Gegenstand des Dokuments: Als Gegenstand wird ein in einem Dokument behandeltes Thema (Sachverhalt) bezeichnet. Ein solcher wird durch einen oder mehrere Begriffe, die in einer sachlichen Beziehung zueinander stehen, repräsentiert (vgl. §§ 7,1; 13,4). Der Dokumenteninhalt besteht aus einem oder mehreren Gegenständen.

Begriff: Ein Begriff ist das Ergebnis einer Abstraktion. In Allgemeinbegriffen (abstrakten Begriffen) sind gemeinsame Merkmale einer Menge individueller Gegenstände materieller und immaterieller Art zusammengefasst. Zu den Begriffen zählen auch Individualbegriffe, sie werden in der Regel durch Individualnamen wiedergegeben.

Bezeichnung (Benennung): Eine Bezeichnung ist ein sprachlicher Ausdruck für einen Allgemeinbegriff oder einen Individualbegriff.

Schlagwort (Deskriptor, Vorzugsbezeichnung, Ansetzungsform): Ein Schlagwort ist eine terminologisch kontrollierte Bezeichnung, die in Indexierung und Retrieval für einen Begriff aus einem Dokumenteninhalt verwendet wird.

Relation (Beziehung): Eine Relation ist die Beziehung zwischen mehreren Bezeichnungen, die denselben Begriff repräsentieren (Äquivalenzrelation), oder die Beziehung zwischen verschiedenen Begriffen aufgrund ihrer gemeinsamen Merkmale (Begriffsbeziehung: hierarchische und assoziative Relation)

Verweisung: Eine Verweisung ist die Darstellung der Relation zwischen Bezeichnungen bzw. Begriffen mit Hilfe einer abgekürzten Verweisformel (z.B. den Symbolen BS, BF usw.). Verweisungen dienen zur Führung des Benutzers, vor allem von nicht zugelassenen auf die zugelassene Bezeichnung (Vorzugsbezeichnung) oder auf hierarchisch übergeordnete, untergeordnete oder auf verwandte Begriffe (vgl. §12).

Die Regeln für den Schlagwortkatalog (RSWK)

Aktuell gültige 3. Auflage:

Regeln für den Schlagwortkatalog: RSWK / erarb. Von der Expertengruppe RSWK auf der Grundlage der von der Kommission des Deutschen Bibliotheksinstituts für Sacherschließung bearb. 2. Aufl. (1991).
Grundwerk : 3, überarb. und erw. Aufl., - Berlin: Dt. Bibliotheksinst., 1998. - ISBN 3-87068-591-3
1. Ergänzungslieferung 2001, 2. Ergänzungslieferung Herbst 2002
Das Regelwerk findet man auf dem Server des Ehemaligen Deutschen Bibliotheksinstituts (EDBI) unter Publikationen: http://www.dbi-berlin.de/dbi_pub/einzelpu/regelw/rswk/rswk_00.htm

Regelwerksänderungen wurden bis 2001 im "Bibliotheksdienst" als **RSWK-Mitteilungen** zur Diskussion gestellt und ebendort nach einer gewissen Einspruchsfrist als verbindliche Änderungen veröffentlicht (s. a. http://bibliotheksdienst.zlb.de/2001/01_02_07.htm, RSWK 3. Aufl., dritte Mitteilung). Seitdem hat Die Deutsche Bibliothek (Arbeitsstelle für Standardisierung) die Publikation von Änderungen übernommen.

Charakteristika der RSWK

1. Das Regelwerk ist **universell**, d.h. für alle Wissenschaftsdisziplinen und für Bibliotheken unterschiedlicher Aufgabenstellungen gedacht (z.B. wissenschaftliche u. öffentliche Bibliotheken).
2. Die Regeln versuchen, **möglichst ausführlich und präzise** zu sein, damit sie sowohl für zentrale Dienstleistungen (z.B. Deutsche Bibliothek) wie auch für kooperative Sacherschließung (z.B. Bibliotheksverbände) verwendet werden können. Deshalb **stützen sie sich in starkem Maße auf Normierung durch Listen** (Anhänge der RSWK, besonders die Formschlagwortliste, Nachschlagewerke und Schlagwortnormdatei (SWD)).
3. Das Regelwerk ist für **EDV-gestützte Kataloge geeignet**, auch für OPACs. Seit der 3. Aufl. wird der Online-Katalog als Standard-Katalogform vorausgesetzt.
4. Wahl des **spezifischen** (engen) Schlagworts.
5. Beibehaltung der **herkömmlichen Kategorieneinteilung** (Personen, Geographika, Sachbegriffe, Zeit und Form), die durch Indikatoren in der SWD ergänzt wird.
6. Prinzip der **Kettenbildung** von Schlagwörtern, dabei Gleichrangigkeit von Geographikum und Sachbegriff. Die Kette soll den Inhalt des Dokuments möglichst präzise erfassen (spezifisches Schlagwort).
7. Orientierung **am Sprachgebrauch** bei der Schlagwortbildung; Maßstab hierfür sind allgemeinverbindliche Nachschlagewerke (vor allem die großen deutschen Enzyklopädien, dann deutschsprachige Fachnachschlagewerke).
8. Die RSWK **orientieren sich bei den Personennamen, Körperschaften und Geographika an den Ansetzungsregeln von RAK**, mit der wichtigen **Ausnahme**, dass sie in allen 3 Fällen **dem deutschen Sprachgebrauch den Vorzug** geben.
9. Die Regeln bieten **möglichst viele Sucheinstiege und Verweisungen**, wobei auch Listenkataloge und Listenfunktionen in OPACs berücksichtigt werden.

Anwendung der RSWK

Die RSWK-Anwendung ist optimiert für das heute gängige Umfeld: kooperative Sacherschließung in Verbänden und Fremdatennutzung als tragende Säule, daneben Anwendung in einzelnen Bibliotheken und zunehmend auch für Datenbanken anderer Bereiche (z.B. im Buchhandel). Die SWD wird darüber hinaus im

Arbeitshilfen und Arbeitsmaterialien zu den RSWK

Als Arbeitsmaterialien dienen neben dem **Regelwerk** die **Schlagwortnormdatei (SWD)**, die **Liste der fachlichen Nachschlagewerke zu den Normdateien**, die **Praxisregeln zu den RSWK und zur SWD** und die **Beispielsammlung zu den RSWK**.

Die Schlagwortnormdatei (SWD)

Sie dient der **terminologischen Kontrolle** und hat bei der Wahl des Schlagworts **erste Priorität**. Ende 2001 enthielt sie über **500 000 Schlagwortsätze** (davon 150 000 Sachschlagwörter) und **630 000 Verweisungen**, dazu zahlreiche Verwendungshinweise, die manchen Blick ins Regelwerk ersparen.

Die SWD erscheint halbjährlich auf Mikrofiche und zusammen mit der PND und GKD auf der Normdaten-CD-ROM Der DB sowie auf Magnetband und Diskette mit wöchentlichem Änderungsdienst. Im Internet findet man sie online auf dem Server des HBZ :
<http://apollon.hbz-nrw.de:4505/ALEPH/ANQJI57R2D232IRY6LJ6BI2VV/QF6CYNNA2Y3X778C.JD26GKEQ1-02312/start/hbz12>

oder im Rahmen der Katalogdatenbank ILTIS Der Deutschen Bibliothek
http://z3950gw.dbf.ddb.de/z3950/zfo_get_file.cgi?fileName=DDB/searchForm.html

Sie wird als **Online-Datei** bei der **Deutschen Bibliothek** geführt, die auch verantwortlich zeichnet. Neben den Mitarbeitern der Deutschen Bibliothek bringen Vertreter aller deutscher Bibliotheksverbände (GBV, KOBV, HBZ, HEBIS; SWB, BVB) und des Verbunds wissenschaftlicher Bibliotheken Österreichs, der Schweizerischen Landesbibliothek (SLB) sowie des Kunstbibliotheken-Fachverbundes Florenz, München, Rom und des Verbundes der kirchlich-wissenschaftlichen Bibliotheken online täglich Termini ein, die sich aus der Beschlagwortung ihrer Neuzugänge ergeben. In jedem Verbund gibt es eine bis maximal zwei **Redaktionen**, die für die von ihm eingebrachten Ansetzungen verantwortlich sind. Die Partner überprüfen diese Ansetzungen gegenseitig, was durch Online-Abwurf der Neuansetzungen bzw. anhand abonnierbarer Änderungsdienste möglich ist. **Änderungs- und Korrekturwünsche** werden online ausgetauscht, d.h. an den betreffenden Schlagwortsatz wird eine Mitteilung angehängt, die der Adressat abrufen und ggf. bearbeitet.

Die Darstellung der einzelnen Begriffe in der SWD erfolgt nach der **DIN-Norm 1463 für Thesauri**, wie denn überhaupt die SWD sich in Richtung eines Universal-Thesaurus bewegt, ohne dieses Ziel je völlig zu erreichen, da ihr Begriffsschatz offen ist und von den Dokumenten abhängt, die bisher zur Erschließung vorgelegen haben. Unter jeder Ansetzungsform sind sämtliche dazugehörigen Schlagworteintragungen zu einem Satz zusammengeführt; wichtig für den Anwender sind dabei vor allem auch Quellenangaben (Q), Definitionen (D) bzw. Verwendungshinweise sowie die Angabe der Relationen, in denen ein Begriff zu einem anderen steht, d.h. ob es übergeordnete (OB), untergeordnete (UB) oder verwandte (VB) Termini gibt.

Identifikations-Nr.	4037589-4 s	Indikator
Ansetzungsform	Marketing Q M; SYS 10.11e BF Absatzpolitik Absatzwirtschaft . . VB Absatz UB Beschaffungsmarketing Direktmarketing	Benutzungshilfen Verweisungsformen

Die Liste der fachlichen Nachschlagewerke zu den Normdateien

Bei der Ansetzung stützen sich die RSWK in hohem Maß auf allgemeine und fachliche Nachschlagewerke. In der Ausgabe vom April 2001 enthält die Liste der fachlichen Nachschlagewerke 1281 überwiegend deutschsprachige Titel. Der Hauptteil der Liste ist nach den 65 Sachgruppen der deutschen Nationalbibliographie sowie nach Formalgruppen für Personennamen (p), Werktiteln (t), Geographika und Ethnographika (g) und Körperschaften (k) gegliedert. Die hier aufgeführten Werke werden für die **Terminologiekontrolle neu anzusetzender Begriffe vordringlich herangezogen** und in den Schlagwortsätzen als Quellenangaben aufgeführt. Auf den Hauptteil folgt eine als "Rangfolge der Nachschlagewerke" betitelte Liste, der man entnehmen kann, in welcher Reihenfolge die aufgeführten Nachschlagewerke heranzuziehen sind. Die Liste der fachlichen Nachschlagewerke wird ständig aktualisiert, erscheint als Papierausgabe und ist auf dem Informationsfiche der Mikroficheausgabe der SWD zu finden. Online kann man die Titel unter <http://support.ddb.de/swd/listeNSW/index.htm> einsehen und die Rangfolge unter <http://www.bibliothek.uni-augsburg.de/allg/swk/nsw.html>.

Die Praxisregeln zu den RSWK und zur SWD

2000 erschien das erstmals 1993 veröffentlichte Grundwerk der Loseblattausgabe in der 3.Aufl. unter dem Titel

Praxisregeln zu den Regeln für den Schlagwortkatalog (RSWK) und der Schlagwortnormdatei (SWD) / Die Deutsche Bibliothek. Red.: Barbara Kelm; Sigrid Bellgardt. - Leipzig ; Frankfurt am Main ; Berlin: Die Dt. Bibliothek. - Losebl.-Ausg. - ISSN 0941-6323

Die RSWK können nur Prinzipien bieten, nach denen bei der Beschlagwortung zu verfahren ist, und die SWD liefert den normierten Wortschatz dazu. Beide können aber nicht auf jedes Einzelproblem eingehen und den sich aus der täglichen Arbeit ergebenden Anforderungen jeweils Rechnung tragen. So bleibt ein Spielraum für unterschiedliche Regelauslegungen, oder es treten neue Sachverhalte auf, denen aufgrund von Regelwerk und Normdatei zu entsprechen nicht immer zweifelsfrei möglich ist. So mussten aus der täglichen praktischen Arbeit heraus immer wieder Vereinbarungen getroffen werden, die als Praxisregeln veröffentlicht und jährlich neu ergänzt werden. Sie stellen in erster Linie ein Hilfsmittel und ein Informationsorgan für den Kreis derer dar, die aktiv mit den RSWK arbeiten und an der SWD- Arbeit redaktionell beteiligt sind.

Die Beispielsammlung zu den RSWK

Beispielsammlung zu den Regeln für den Schlagwortkatalog / Deutsches Bibliotheksinstitut. Bearb. von Gerhard Stumpf... – 2., vollst. neu bearb. Aufl. nach der 3. Aufl. des Regelwerks. - Berlin: Dt. Bibliotheksinst., 1999. - 218 S. - ISBN 3-87068-615-4

Hier werden an Hand von 197 kommentierten **Beispielen**, gegliedert nach dem Aufbau der RSWK, in **kurzen erläuternden Texten** die wesentlichen Bestimmungen der einzelnen Regelwerksteile vorgestellt. Das Ganze wird erschlossen durch ein Beispielregister und ein Sachregister.

Kooperation in der verbalen Sacherschließung der deutschen Bibliotheken

In Deutschland gibt es sechs Katalogisierungsverbände, denen die meisten wissenschaftlichen Bibliotheken und zunehmend auch die öffentlichen Bibliotheken angeschlossen sind. Im einzelnen sind dies:

- **der Gemeinsame Bibliotheksverbund mit Niedersachsen, Sachsen-Anhalt, Thüringen, Hamburg und Bremen (GBV)**
- **der Nordrhein-westfälische Verbund (HBZ)**
- **der Hessische Verbund (HeBIS)**
- **der Südwestdeutsche Verbund mit Sachsen (SWB)**
- **der Bibliotheksverbund Bayern (BVB) und**
- **der Berlin-Brandenburger Verbund (KOBV)**

Hinzu kommen der Verbund der österreichischen wissenschaftlichen Bibliotheken und die Bibliotheken der Deutschschweiz als Nutzer von Fremddaten der Deutschen Bibliothek und – von Fall zu Fall – auch anderer Verbünde.

In den Katalogisierungsverbänden erschließen die angeschlossenen Bibliotheken unter Nutzung von Fremddaten ihre Bestände arbeitsteilig. Es versteht sich von selbst, dass eine **kooperative Erschließung** ein **gemeinsames Regelwerk** voraussetzt. Für die **Formalerschließung** sind dies in Deutschland und Österreich die Regeln für den alphabetischen Katalog (**RAK-WB**)

In den **1990er Jahren** haben sich die **RSWK** im RAK-Bereich **weitgehend etabliert**. Alle Verbünde haben sich für die RSWK, entweder mit oder ohne Kettenbildung, und für die Schlagwortnormdatei entschieden und ihre Datenstruktur entsprechend eingerichtet. Das Austauschformat MAB2 umfasst diesen Standard für Titeldaten und SWD. Die Teilnehmerbibliotheken stellen ihre lokale Sacherschließung nach und nach auf die kooperative Sacherschließung nach den RSWK um.

Es gibt viele Gründe, die für ein einheitliches Regelwerk sprechen, ausschlaggebend für die meisten Bibliotheken war jedoch die Möglichkeit der **Fremddatennutzung**. Ist man bei der Formalerschließung noch darauf angewiesen, dass zum Zeitpunkt der Einarbeitung eines Mediums bereits Fremdleistung vorliegt, so sind die meisten Verbünde dazu übergegangen, die Sacherschließungsdaten der Deutschen Bibliothek automatisch nachzuführen, sobald sie vorliegen, und die angeschlossenen Bibliotheken beschlagworten nur noch diejenigen Bücher, bei denen keine Fremdleistung der Deutschen Bibliothek zu erwarten ist.

RSWK und Online-Katalog

Im Herbst 1994 erschienen die **Empfehlungen "Sacherschließung in Online-Katalogen" der Expertengruppe Online-Kataloge** des Deutschen Bibliotheksinstituts.

Sacherschließung in Online-Katalogen / Kommission des Deutschen Bibliotheksinstituts für Erschließung und Katalogmanagement, Expertengruppe Online-Kataloge; Deutsches Bibliotheksinstitut (Hrsg.: Friedrich Geißelmann). - Berlin: Dt. Bibliotheksinstitut 1994., 106 S.

Hierin wurde auf **Anpassungen der RSWK an den Online-Katalog** sowie **mögliche ergänzende Methoden zu den RSWK** eingegangen. Des weiteren findet man hier **zahlreiche Überlegungen zur Gestaltung des OPAC**, die in die Planung jedes OPAC mit einbezogen werden sollten.

Die Expertengruppe hatte zur **Stellungnahme** zu diesen Empfehlungen aufgerufen, die im Frühjahr 1995 auf einem **Kolloquium in Weimar** vorgetragen und diskutiert wurden. Ergebnis des Kolloquiums war eine breite Zustimmung in fast allen wesentlichen Punkten sowie zusätzliche Anregungen für die Neuauflage der RSWK und der Beschluss, die vorgeschlagenen Änderungen in die neue (3.) Auflage der RSWK aufzunehmen.

Wesentliche Punkte der Empfehlungen in Auswahl

- Beschränkung auf notwendige, nachvollziehbare Regeln
- Einheitliche Ansetzung von Personennamen, Körperschaftsnamen und Einheitssachtiteln
- Vermeidung von Doppelerfassung bei Sach- und Formalerschließung
- Parallele verbale und klassifikatorische Sacherschließung
- Integration von Aufbau und Aufgabe der SWD ins Regelwerk
- Einschränkung des Pleonasmusverbots
- Verstärkte Anwendung des Homonymenzusatzes
- Verwendung von Zeit- und Ländercodes
- Abschaffung der Pauschalverweisungen
- Reduzierung der Kompositabildung (stärkere Zerlegung)
- Abschaffung von Ansetzungsketten beim Sachschlagwort
- Plädoyer für Kettenlisten im OPAC als Browsing-Instrument zur Relevanzbeurteilung
- Abbildung der Thesaurusstruktur und der Begriffsrelationen der SWD im OPAC

Modifizierungen und weitere Empfehlungen als Ergebnis des Kolloquiums in Auswahl

- Weitgehende Abschaffung der Regeln zur Kettensyntax
- Permutationsregeln nur noch fakultativ
- Die Erschließung unselbständiger Werke soll weiterhin freigestellt sein
- Erhöhung der Anzahl der Kettenglieder pro Schlagwortkette

Verbale Sacherschließung im Informationsverbund Deutschschweiz (IDS)

Über Entwicklung und aktuellen Stand im IDS kann man sich auf der Homepage des IDS (<http://www.zb3.unizh.ch/ids/>) und auf der Seite „Sachkatalogisierung im Verbund“ (IDS Basel/Bern) http://www.ub.unibas.ch/verbund/ibb/S_Kat/_index.htm informieren.

Library of Congress Subject Headings (LCSH)

Entstehung der LCSH

Grundlage der ersten amerikanischen Schlagwortliste bildeten **Cutters "Rules for a printed dictionary catalogue"**, die 1876 veröffentlicht wurden. Die LoC brachte die erste Ausgabe der LCSH in Lieferungen zwischen 1909 und 1914 heraus.

Im Vergleich zu anderen Ländern hat das **amerikanische Bibliothekswesen sehr früh einheitliche Organisationsformen entwickelt**. Dies konnte geschehen, weil Leistung und Durchsetzungsvermögen einzelner Personen und Institutionen mit einer **besonderen Bereitschaft zusammentrafen, "Fremdleistung"** aus praktischen und ökonomischen Gesichtspunkten **zu übernehmen**. Größte Bedeutung kommt dabei dem Verkauf der **LC-Titelaufnahmen** zu, der 1901 einsetzte und 1973 bereits 25 000 Besteller in den USA und 1 200 Abnehmer in anderen Ländern gefunden hatte.

Die LCSH wirkten **normbildend für amerikanische Spezialschlagwortlisten** und prägten auch die **Schlagwortverzeichnisse anderer Länder**. Sie sind heute in Papierform, als Mikroficheausgabe, auf CD-ROM und als Magnetband verfügbar. Die Papierausgabe gibt es bereits in der 21. Auflage. Sie enthält mehr als 250 000 Headings und References.

Library of Congress / Cataloging Distribution Service (Hrsg.): Library of Congress Subject Headings. 25th. ed.: 5 Vols. 2001.

Ein webbasierter Zugang zu den Subject Headings hat wurde von der Mississippi State University entwickelt: **Library of Congress Subject Headings Database** <http://fantasia.cs.msstate.edu/lcshdb/index.cgi>

Detaillierte Regeln zur Schlagwortbildung und Anwendung der LCSH für den Katalogisierer findet man in:

Library of Congress / Cataloging Distribution Service (Hrsg.): Subject Cataloging Manual: Subject Headings. 5th. ed.: 4 vols. 1996.

Sowohl zu den Subject headings als auch zu den Katalogisierungsregeln gibt es einen Update Service.

Grundprinzipien der LCSH

1. Die LCSH sind **universell**, d.h. sie umfassen alle Wissenschaftsdisziplinen.
2. Die Wahl des Schlagworts orientiert sich am **natürlichen Sprachgebrauch**, selbst wenn man dadurch Inkonsequenzen in Kauf nehmen muss.
3. Wahl des **spezifischen** (engen) Schlagworts.
4. Die Beschlagwortung erfolgt **normiert**. **Synonyme** werden unter einem Heading zusammengeführt und **Homonyme** werden durch Zusätze unterschieden.
5. **Personennamen, Körperschaften** und **Geographika** werden nach den Anglo-American Cataloging Rules (**AACR**) angesetzt und i.d.R. nicht in die Liste aufgenommen.

6. **Vermischung von Prä- und Postkoordination:** Komplexe Sachverhalte werden im LCSH-System entweder durch Prä- oder Postkoordination dargestellt. Die Anwendung der LCSH stellt also eine Mischung von syntaktischer und gleichordnender Indexierung dar, wobei die Präkoordination allerdings bevorzugt wird und auch den Hauptanteil der bibliographischen Einträge ausmacht.

II Maschinelle Indexierung zur verbesserten Literaturschließung in Online-Systemen (MILOS), durchgeführt an der ULB Düsseldorf

Die konventionelle Form der **intellektuellen Erschließung** (z.B. RSWK) ist personell **aufwändig** und dadurch, trotz des kooperativen Effekts, meist beschränkt auf die laufenden Neuzugänge. **Große Mengen von Titeldaten** sind und bleiben daher **ohne Sacherschließung**, was den Zugriff auf die betreffenden Dokumente erheblich erschwert.

In den Jahren 1993-1996 wurden an der ULB Düsseldorf in enger Zusammenarbeit mit der Fachrichtung Informationswissenschaft der Universität des Saarlandes (Prof. Dr. Harald H. Zimmermann) die beiden von der Deutschen Forschungsgemeinschaft geförderten Projekte MILOS I und MILOS II durchgeführt.

Im Rahmen des ersten Projekts ging es um die **Anpassung** des von Zimmermann entwickelten Systems zur automatischen Indexierung, **IDX**, an die besonderen Bedingungen eines **Einsatzes in Bibliotheken**. Dazu wurde im Laufe des Projekts u.a. eine Indexierung mit dem vollständigen maschinenlesbaren Titelbestand der ULB Düsseldorf durchgeführt. Auf dieser Basis wurden die für die Arbeit des rein wörterbuchbasierten Systems IDX entscheidenden elektronischen Wörterbücher erheblich erweitert. Schwerpunkt des **Nachfolgeprojekts MILOS II** war die **Einbeziehung semantischer Komponenten** in die automatische Indexierung, wobei die Integration der **Schlagwortnormdatei** in das System zentraler Bestandteil war.

Das Indexierungssystem IDX hat in der eingesetzten und speziell angepaßten Version OPAC-IDX folgenden Funktionsumfang:

1. **Grundformermittlung**, d.h. Ermittlung von Grundformen zu Wortformen (Pluralformen, flektierte Formen)
2. **Markierung bzw. Eliminierung von Stopwörtern**
3. **Dekomposition und Derivation:** Komposita werden wörterbuchgestützt in ihre sinnvollen Wortbestandteile zerlegt. Derivationen, d.h. Wortableitungen werden in Richtung Verb – Substantiv und Adjektiv -- Substantiv durchgeführt.
4. **Wortrelationierung:** Unter dem Begriff Wortrelationierung wird eine Reihe von Operationen subsummiert, die durch das Relationenkonzept der verwendeten Wörterbücher unterstützt werden.
5. **Mehrworterkennung und Wortbindestrichergängung:** Die Erkennung von Mehrwortbegriffen erleichtert das spätere Retrieval feststehender Wendungen. Die im Deutschen sehr beliebte Tilgung von Wortbestandteilen führt bei Stichwortindexierung zu ernsthaften Problemen, denen durch die Wortbindestrichergängung begegnet wird.
6. **Wortbezogene Übersetzung** durch Zuschaltung entsprechender Übersetzungswörterbücher. Die in IDX integrierte Übersetzungsfunktion leistet keine vollständige Übersetzung von Titeln, sondern beschränkt sich auf die wortweise Übersetzung relevanter Begriffe.

Im Rahmen von MILOS I standen die grammatikalischen Funktionen 1. bis 3. im Vordergrund, semantische Relationierungen (4.) wurden lediglich teilweise in die Indexierung einbezogen und bildeten dann den Schwerpunkt von MILOS II. Die wortbezogene Übersetzung wurde nach ersten Versuchen zunächst zurückgestellt und danach gesondert im Rahmen eines von der Europäischen Union geförderten Projekts, CANAL/LS (Catalog with natural language access / linguistic server) weiterentwickelt.

Im Projekt wurden **alle für die sachliche Recherche relevanten Kategorien der Titelaufnahmen** (Hauptsachtitel, Zusatz zum Sachtitel, lokale Schlagwörter u.a.) **in mehreren Schritten indexiert**. Der eigentlichen Indexierung geht eine automatische **Rechtschreibkontrolle** voraus, die sowohl fehlerhafte als auch in den Wörterbüchern nicht vorhandene Begriffe erkennt. Diese werden anschließend in den **Wörterbüchern geändert bzw. ganz neu eingetragen**.

Für jede der drei Indexierungssprachen (Deutsch, Französisch, Englisch) gibt es ein sogenanntes **Identifikationswörterbuch**, das **alle Grundformen und möglichen Endungen** enthält, sowie ein sogenanntes **Relationenwörterbuch**, über das **Dekomposition und Derivation** gesteuert werden können. Hinzu kommen für die Verwendung mit englischen und französischen Titeln die **Übersetzungswörterbücher** Englisch -- Deutsch und Französisch -- Deutsch.

Ergebnis

Beim Einsatz von MILOS in einem Online Public Access Catalogue (OPAC) ergeben sich folgende Funktionserweiterungen:

- Retrieval mit Grundformen, sinntragenden Teilwörtern, Synonymen und Derivationen, d.h. auch das Trunkierungsproblem entfällt weitgehend,
- Retrieval mit sinntragenden Mehrwortbegriffen, d.h. das Adjacency-Problem entfällt weitgehend
- Retrieval unter Einbeziehung von Übersetzungsäquivalenten

Bei allen bisherigen Tests, die mit dem Kernsystem IDX im Rahmen des MILOS-Projekts durchgeführt wurden, bestätigten eine deutliche Verbesserung gegenüber der reinen Stichwortsuche.

Zur Zeit hat das System noch folgende Schwachstellen:

- Es erfolgt keine semantische Differenzierung der Grundformen. Z. B. wird nicht zwischen Schloss (Verschluss) und Schloss (Gebäude) unterschieden.
- Grundformermittlung und Wortzerlegung kann zur Überindexierung führen. Z. B. Abbildung von "weinen" (Verb) auf "Wein" (Substantiv) und Auffassung von Eigennamen "Goldstein" als Komposita.

Dokumente zu MILOS findet man unter http://www.uni-duesseldorf.de/WWW/ulb/mil_home.htm

Folgerungen

Automatische Indexierung kann insbesondere die Verfügbarkeit der im allgemeinen **nicht verbal erschlossenen konvertierten Altdaten verbessern**. Die Ausnutzung und Auswertung weiterer Erschließungsressourcen wie Abstracts, Inhaltsverzeichnisse und/oder Sachregister könnte dabei zu einem noch besseren Ergebnis beitragen. Daher hat die **UB Düsseldorf von April 1997 bis März 1999 ein weiteres Projekt "KASCADE"** durchgeführt, das die o. g. zusätzlichen inhaltsrelevanten Informationen im Scanning-Verfahren ergänzt und die angereicherten Titeldaten aufbauend auf MILOS automatisch erschließt. Dokumente zu KASCADE findet man unter http://www.rz.uni-duesseldorf.de/ulb/kas_home.htm.

Die **Zukunft der Inhaltserschließung** mit Blick auf den Online-Katalog ist am ehesten in einer **Kombination von automatischen und intellektuellen Verfahren** zu suchen. Die nur begrenzt verfügbaren Ressourcen zur intellektuellen Beschlagwortung könnten gezielt für Titel eingesetzt werden, die für eine automatische Verarbeitung ungeeignet sind; Doppelarbeit in der Erschließung, wie die Beschlagwortung mit Titelstichwörtern, könnte entfallen.

Empfehlenswerte Linksammlung:

„Internet-Informationen zur Sacherschließung“ der UB Augsburg:
http://www.bibliothek.uni-augsburg.de/allg/swk/sacher_allg.html