

Ulrike Junger

Überlegungen zur Weiterentwicklung der Sacherschließung im Gemeinsamen Bibliotheksverbund/GBV*

Der gegenwärtige Stand der Sacherschließung in den Datenbanken des GBV wird beleuchtet. Obwohl es seit einigen Jahren ein Konzept zur kooperativen Sacherschließung gibt, ist dennoch ein großer Teil der Daten über eine inhaltliche Recherche nicht greifbar. Mehrere Möglichkeiten werden vorgestellt, um hier eine Verbesserung zu erreichen. Dazu gehören Konkordanzen zwischen verschiedenen bibliothekarischen Klassifikationen, der Einsatz von Verfahren zur automatischen verbalen Indexierung, die Verknüpfung von verbalen und klassifikatorischen Sacherschließungselementen sowie verbesserte Suchoberflächen. Einige dieser Maßnahmen sollen im GBV in der nächsten Zeit umgesetzt werden.

Considerations for the further development of subject indexing within the Gemeinsamer Bibliotheksverbund/GBV (North German Library Network)

The present situation in terms of subject indexing for the GBV databases is examined. Although the participants in the library network have agreed on guidelines for shared subject indexing, a large part of the data cannot be reached through subject searching. Various possibilities to improve this situation are presented. They comprise crosswalks between classifications, use of methods for automatic verbal indexing, linkings between verbal and classificatory elements and improved user interfaces. The GBV will implement some of these measures in the near future.

Considérations pour le développement ultérieur de l'indexation des matières au Gemeinsamer Bibliotheksverbund/GBV (Réseau Bibliothécaire de l'Allemagne du Nord)

La situation actuelle de l'indexation des matières dans les bases de données du GBV est examinée. Malgré que les bibliothèques participantes au GBV s'aient entendu sur des règles de l'indexation des matières communes, une large part des données n'est pas accessible par des recherches à matière. Plusieurs possibilités pour améliorer la situation sont présentées. Ils comprennent des concordances entre des classifications différentes, l'usage des méthodes de l'indexation automatique, les jonctions entre des éléments verbaux et systématiques et des interfaces utilisateur améliorés. Le GBV va à réaliser quelques-unes de ces mesures dans un avenir proche.

Eine Vorbemerkung sei den folgenden Ausführungen vorangestellt: *Sacherschließung ist kein Selbstzweck.* Sie dient nicht der Erfüllung von Regelwerken und Richtlinien, sondern muß wie alle bibliothekarischen Dienstleistungen auf den Endbenutzer ausgerichtet sein. An diesem Grundsatz müssen sich alle Überlegungen orientieren.

Gegenwärtiger Stand der Sacherschließung im GBV

1. *Ein großer Teil der Daten in den Datenbanken des GBV kann über eine sachliche Suche nicht erreicht werden, da klassifikatorische und/oder verbale Erschließung fehlt oder nicht suchbar ist. Dies trifft bei der Katalogdatenbank insbesondere für Alt- und Retrodaten, aber auch für ausländische und graue Literatur zu. Für die Datenbanken, die unselbständige Literatur verzeichnen, wird überhaupt keine Sacherschließung betrieben.*

Bibliothekarische Sacherschließung, und hier besonders verbale Erschließung wird derzeit in der Regel nur für Monographien betrieben. Im GBV sieht die Situation am besten für neuere deutschsprachige Titel aus, da hier die Schlagwortketten der Deutschen Bibliothek als Fremdleistung übernommen werden können. Für die in der Katalogdatenbank nachgewiesene ausländische

und graue Literatur sowie für Altdaten jedoch fehlt eine einheitliche, durchgängige Erschließung.

Noch unbefriedigender ist die Situation für unselbständige Literatur. Um deren Erschließung kümmern sich die Bibliotheken nur noch in Ausnahmefällen, wie etwa für Regionalbibliographien. Diese Literatur, die in der Regel aber den aktuellen Stand der Forschung wiedergibt, ist für den Benutzer wegen fehlender Sacherschließung über inhaltliche Recherchen nicht greifbar. Es bleibt zu fragen, ob es sinnvoll ist, daß Bibliotheken und Bibliotheksverbände ihren Benutzern zwar vermehrt solche Datenbanken anbieten, jedoch keinen umfassenden sachlichen Zugriff darauf ermöglichen. Im GBV betrifft das beispielsweise die Datenbanken Internationale Bibliographie der geistes- und sozialwissenschaftlichen Zeitschriftenliteratur/IBZ, Online Contents/OLC und den Periodicals Contents Index/PCI.

2. *Sacherschließungsdaten rühren aus einer Reihe unterschiedlicher Klassifikationen und verbaler Erschließungssysteme her, die zudem noch in unterschiedlicher zeitlicher Schichtung vorliegen. Dem Benutzer wird in der Regel nur eine Auswahl davon angeboten.*

* Überarbeitete Version eines kurzen Vortrags für den Workshop „Perspektiven der Sacherschließung im GBV“, der am 12.3.1998 in Hannover stattfand.

Vor der Durchführung einer sachlichen Recherche müßte sich der Benutzer derzeit über folgende Tatsachen bewußt sein: Erstens gibt es Nachweise ohne sachliche Erschließung. Zweitens sind die vorhandenen Sacherschließungsdaten heterogen, das heißt unterschiedliche Systematiken, Schlagwortsysteme und Regelwerke liegen ihnen zugrunde. Beides unterliegt darüberhinaus noch einer zeitlichen Schichtung. Dies ist für den Benutzer bei einer thematischen Recherche nicht durchschaubar. Daher wird ihm in der Regel im OPAC nur eine Auswahl an Sacherschließungsdaten angeboten. Das bedeutet aber, daß ein großer Teil der doch vorhandenen Sacherschließungsdaten für den Benutzer keinen Nutzen erbringt. Als Beispiel sei hier nur genannt, daß in der Katalogdatenbank des GBV zwar in großem Umfang Schlagwörter der Library of Congress, der National Library of Medicine und der British Library enthalten sind, aus den genannten Gründen bisher aber auf deren Indexierung verzichtet wurde.

3. Die Beteiligung an der kooperativen Sacherschließung im GBV ist vor allem im Hinblick auf die verbale Erschließung nach RSWK/SWD unzureichend.

Im Zusammenhang mit der Einführung des PICA-Bibliothekssystems wurde im damaligen Bibliotheksverbund Niedersachsen/Sachsen-Anhalt auch ein Konzept zur gemeinsamen Sacherschließung entworfen und in den „Richtlinien zur kooperativen Sacherschließung“ niedergelegt. Es umfaßt eine klassifikatorische (die Erschließung mit der Basisklassifikation) und eine verbale Komponente (die Erschließung nach etwas modifizierten RSWK mit der Schlagwortnormdatei als verbindlichem Thesaurus), dazu Erfassungsmöglichkeiten für Zeit- und Forms Schlagwörter. Dafür stehen verschiedene Felder auf zentraler Ebene der Katalogdatenbank zur Verfügung.

Ab Januar 1993 wurden die Richtlinien für die kooperative Sacherschließung im Bibliotheksverbund Niedersachsen/Sachsen-Anhalt eingeführt, und auch alle hinzukommenden Verbundpartner haben im Prinzip dieses Konzept übernommen.

Im Frühjahr 1997 führte die Zentralredaktion Sacherschließung unter den Teilnehmerbibliotheken des GBV eine Umfrage zur Beteiligung an der kooperativen Sacherschließung durch. Als wichtigste Resultate seien hier nur genannt, daß von 58 Bibliotheken aller Sparten, deren Fragebögen ausgewertet wurden, die eine Hälfte (genauer 30 Bibliotheken) die Basisklassifikation für alle Neuerwerbungen anwenden, die andere Hälfte (28 Bibliotheken) dagegen überhaupt nicht. An der kooperativen verbalen Erschließung mit der Schlagwortnormdatei beteiligen sich 31 von 58 Bibliotheken überhaupt nicht, 21 beschlagworten alle Neuerwerbungen, 6 Bibliotheken beschlagworten in Auswahl. Das Ergebnis hinsichtlich der Beteiligung an der verbalen Erschließung ist besonders bedauerlich, da eine Reihe von Studien gezeigt hat, daß Benutzer verbale Sucheinstiege bevorzugen und mit der klassischen systematischen Suche (d.h. der Eingabe einer Notation) schlecht zurechtkommen. Insgesamt muß die Beteiligung an der kooperativen Sacherschließung im GBV als nicht zufriedenstellend bezeichnet werden.

Welche Gründe wurden von den befragten Bibliotheken für die mangelnde Beteiligung an der verbalen Sacherschließung genannt?

- Der Aufwand für die verbale Sacherschließung wird oft als zu hoch empfunden, das Regelwerk und insbesondere seine Vorschriften zur Neuansetzung von Schlagwörtern gelten als zu kompliziert.
- Die Schlagwortansetzungen der Schlagwortnormdatei werden oft als unbefriedigend empfunden, insbesondere im technisch-naturwissenschaftlichen Bereich. Bemängelt wird einerseits die Verwendung veralteter Terminologie, andererseits die Zerlegung von Begriffen, die in *inhaltlicher Hinsicht* und in *Hinsicht auf das, was der Benutzer vermutlich als Suchbegriff eingibt*, als unsachgemäß empfunden wird. Als Beispiel für den letztgenannten Punkt sei die Zerlegung von Legierungen und Werkstoffe erwähnt: Ist die praktizierte Zerlegung in Einzelbestandteile sachlich angemessen? Durchschauen die Benutzer wirklich dieses Prinzip?
- Die Sacherschließer halten eine gesonderte verbale Erschließung oft für überflüssig, da die Schlagwörter den Titelstichwörtern entsprechen würden bzw. der Titel bereits aussagekräftig genug sei.
- Für einige Bibliotheken ist die Beschlagwortung auf zentraler Ebene nicht differenziert genug, oder ihre Benutzer sind an Fachthesauri und lokale Schlagwortlisten gewöhnt. Dies ist einer der Gründe, warum 17 von 58 Bibliotheken in der Umfrage zur kooperativen Sacherschließung angaben, auf lokaler Ebene zu beschlagworten.

Hier besteht Handlungsbedarf:

- Es wird zu prüfen sein, ob die bevorstehende 3. Auflage der Regeln für den Schlagwortkatalog/RSWK in der praktischen Anwendung den Gegebenheiten von automatisierten bibliothekarischen Informationssystemen wirklich Rechnung trägt, und welcher Aufwand für verbale Inhaltsererschließung und Schlagwortansetzung entsteht (auch verglichen mit der jetzigen Situation). Bei negativer Bewertung müßte eine Neukonzeption der verbalen Sacherschließung in Angriff genommen werden, die eine ernsthafte Entrümpelung der Vorschriften, sinkenden Aufwand für die verbale Inhaltsererschließung und damit eine weitere Verbreitung zum Ziel hat. Die Belange des Zettelkataloges müßten endlich konsequent verabschiedet werden, statt dessen die Erschließung unterschiedlicher Dokumentarten stärker berücksichtigt werden. Intellektuelle Erschließung sollte nur noch dann gemacht werden, wenn ein echter Informationsgewinn dabei entsteht.
- Im Rahmen der kooperativen Sacherschließung im GBV sollte die Vorstellung dessen, was unter verbaler Erschließung zu verstehen ist, erweitert werden. Verbale Erschließung darf nicht auf normiertes Vokabular beschränkt bleiben, vielmehr müssen aus den Titelstichwörtern gewonnene Indexbegriffe und nicht normierte Schlagwörter einbezogen werden, um einen Bereich zwischen normiertem Vokabular und Titelstichwörtern zu schaffen. Auf diese Weise könnten auch die jetzt in lokalen Kategorien abgelegten Schlagwörter verbundweit nutzbar gemacht werden.

4. Die derzeit eingesetzten Indexierungstechniken und Suchoberflächen sind ungenügend.

Die Benutzerforschung hat -grob zusammengefaßt- ergeben, daß Benutzer klassifikatorische Erschließung ignorieren, wenn sie mit der Eingabe von Notationen

verbunden ist. Sie machen kaum aktiven Gebrauch von Thesauri, d.h., sie suchen nicht zuerst die Schlagwortnormdatei nach passenden Suchbegriffen ab. Sie lesen keine Hilfetexte, nutzen Trunkierungsmöglichkeiten und Boole'sche Verknüpfungen gar nicht oder falsch. Sie wissen nichts von bibliothekarischen Regelwerken und können mit bibliothekarischer Begrifflichkeit nichts anfangen. Als Beispiel sei hier nur genannt, daß vielen Benutzern der Unterschied zwischen Stichwort und Schlagwort nicht bekannt ist.

Hinzu kommt, daß die derzeit genutzten Indexierungs- und Suchtechniken verbesserungsbedürftig sind. Es fehlt an Robustheit gegenüber Eingabefehlern, es gibt keine phonetische Indexierung (z.B. Telefon – Telephon) und keine morphologische Bearbeitung der Sucheingaben (z.B. Reduktion von Pluralformen auf den Singular).

Ziele für eine Weiterentwicklung der kooperativen Inhalterschließung im GBV

Die im folgenden genannten Ziele hängen eng zusammen und sollten auch gemeinsam realisiert werden.

- *Es muß eine verbesserte sachliche Erschließung bei höchstens gleichbleibendem Aufwand erreicht werden.*
- *Künftig soll der größte Teil der GBV-Daten über eine thematische Recherche aufzufinden sein, die über die reine Titelstichwortsuche hinausgeht.*
- *Eine verbesserte Nutzung der vorhandenen Sacherschließungselemente, insbesondere aus den Fremddaten, muß erreicht werden.*
- *Ein einheitlicher sachlich-thematischer Zugriff für alle angebotenen Datenbanken im GBV wird angestrebt, d.h. Sacherschließungsdaten unterschiedlicher Provenienz sollen integriert werden.*
- *Es wird ein Maximum an Benutzerfreundlichkeit im Hinblick auf die Suchoberfläche angestrebt. Dies beinhaltet den konsequenten verbalen/natürlichsprachigen Einstieg und die Fähigkeit des Systems, Benutzerfehler abzufangen.*

Denkbare Maßnahmen zur Verbesserung der inhaltlichen Erschließung der GBV-Datenbanken

Für den Bereich der *intellektuellen Erschließung* wird folgendes vorgeschlagen:

- Die RSWK und die Schlagwortnormdatei müssen weiterentwickelt werden: die Komplexität und damit der zu leistende Aufwand für verbale Sacherschließung müssen reduziert werden, um eine weitere Verbreitung zu erreichen.
- Alle denkbaren verbalen Indexate müssen in ein Konzept integriert werden, d.h. durch automatische Verfahren gewonnene Indexate aus Titelstichwörtern, freie Begriffe/Schlagwörter und normierte Schlagwörter.
- Verbale und klassifikatorische Sacherschließungselemente müssen integriert werden. Als Beispiel sei hier ein Projekt an der SUB Göttingen genannt: Dort werden im Rahmen der Bearbeitung der Systematik alten Göttinger Bandkataloges die Notationen mit Begriffen aus der Schlagwortnormdatei versehen.
- Mehrsprachige verbale Erschließung sollte zugelassen werden.

Hinsichtlich der *maschinell gestützten Erschließung* sollte folgendes getan werden:

- Verbesserte Indexierungstechniken müssen eingesetzt werden. Für die Erzeugung zusätzlicher Indexbegriffe durch morphologische Reduktion sollte der Einsatz von Verfahren zur automatischen Indexierung, z.B. MILOS, geprüft werden. Darüber hinaus müssen multilinguale und datenbankübergreifende Indices aufgebaut werden.
- Konkordanzen zwischen verschiedenen Systematiken und der Basisklassifikation müssen eingesetzt werden, um für einen möglichst großen Anteil der GBV-Daten eine einheitliche systematische Zuordnung zu erzielen. Konkret ist dabei bisher an Konkordanzen zwischen der Library of Congress Classification, den Sachgruppen der Deutschen Nationalbibliographie, der Systematik des Göttinger Bandkataloges und der Basisklassifikation gedacht.
- Über Tabellen soll eine maschinelle Verknüpfung von klassifikatorischen und verbalen Sacherschließungsdaten unterschiedlicher Provenienz erreicht werden. Hierbei wäre beispielsweise an Mappings zwischen Notationen der Basisklassifikation und Schlagwörtern oder zwischen Schlagwörtern aus der Schlagwortnormdatei und Schlagwörtern der Library of Congress zu denken. Solche Tabellen bilden auch die Basis für Prognosesysteme, die für die maschinelle Indexierung von unselbständiger Literatur eingesetzt werden könnten (automatische Notationsvergabe).
- Verbesserte Suchmaschinen müssen zum Einsatz kommen, die phonetisches Suchen und die Anwendung von Nachbarschaftsoperatoren ermöglichen sowie ein Relevance Ranking bei der Ergebnisanzeige.

Erste Schritte der Umsetzung

Eine Qualitätsverbesserung in den Datenbanken des GBV hinsichtlich der sachlichen Erschließung läßt sich nur schrittweise erreichen. Zu den Maßnahmen, die ab sofort bzw. im Laufe des Jahres 1998 realisiert werden können, gehören:

Intensivierung der intellektuellen Erschließung

Die Intensivierung und Verbesserung der intellektuellen Erschließung, also die Beteiligung an der kooperativen Sacherschließung ist in erster Linie eine Aufgabe der Verbundbibliotheken. Die so produzierten Sacherschließungsdaten sind ein unabdingbarer Bestandteil für eine Wissensbasis, auf der Computer Aided Indexing und prognostische Verfahren aufsetzen können. Die Zentralredaktion Sacherschließung des GBV wird die Bibliotheken dabei insbesondere im Bereich der verbalen Sacherschließung unterstützen.

Von Seiten der Verbundzentrale ist geplant:

Erstellung und Einspielung von Konkordanzen zwischen der Basisklassifikation und anderen Systematiken

Es ist geplant bzw. teilweise bereits realisiert, Konkordanzen zwischen der Basisklassifikation und anderen Klassifikationen zu erstellen. Auf diese Weise kann nachträglich ein großer Teil der GBV-Daten mit Notationen aus der Basisklassifikation versehen werden. Eine Konkordanz zwischen der Library of Congress Classification und der Basisklassifikation wird voraussichtlich von PICA Leiden zur Verfügung gestellt werden. Damit

könnten rückwirkend diejenigen der 4.8 Mio. Titelaufnahmen der Library of Congress ab 1968 erschlossen werden, die noch keine Notation der Basisklassifikation haben. Eine Konkordanz zu den Sachgruppen der Deutschen Bücherei bzw. der Deutschen Nationalbibliographie wird von der Zentralredaktion Sacherschließung des GBV erstellt. Damit könnte der größte Teil der von der Deutschen Bibliothek erstellten Titelaufnahmen seit 1945 wenigstens grob sachlich zugeordnet werden. Im Rahmen eines bibliotheksinternen Projektes wird derzeit an der SUB Göttingen eine Konkordanz zur Systematik des Göttinger Bandkataloges erarbeitet. Damit könnten mehrere hunderttausend Titelaufnahmen von Altbestand notiert werden.

Einsatz einer neuen Indexmaschine

Noch in der ersten Jahreshälfte 1998 soll eine neue Indexmaschine zum Einsatz kommen, die eine Reihe neuer Möglichkeiten bietet. Dazu gehören die Proximity-Suche sowie ein Relevance Ranking für die Anzeige der Suchergebnisse. Nach der Einführung kann auch mit der Erstellung datenbankübergreifender Indices begonnen werden.

Nachführung der Sacherschließungsdaten der Library of Congress

Im Zuge der Neueinspielung der Daten der Library of Congress werden auch alle Sacherschließungsdaten nachgeführt und für die Indexierung vorbereitet.

Einspielung der Medical Subject Headings als zentrale Normdatei

Die Medical Subject Headings/MESH werden als Normdatei auf zentraler Ebene eingespielt (analog zur Schlagwortnormdatei), um einem langgehegten Wunsch der medizinischen Bibliotheken im GBV Rechnung zu tragen.

Mitarbeit an einem PICA-Projekt zur integrierten thematischen Recherche

Bei PICA in Leiden wird derzeit in mehreren Schritten ein System entwickelt, das im wesentlichen aus zwei Komponenten besteht: einer Wissensbasis, in die klassifikatorische und verbale Sacherschließungselemente verschiedener Herkunft miteinander in Beziehung gesetzt werden sowie einer Benutzerschnittstelle, die natürlichsprachliche Sucheingaben verarbeiten kann. Dies beinhaltet auch den Einsatz computerlinguistischer Methoden. Der GBV plant, diese Software zu übernehmen und für den eigenen Einsatz zu adaptieren.

Anschrift der Autorin:

Ulrike Junger
Niedersächsische Staats-
und Universitätsbibliothek
D-37070 Göttingen
E-Mail: junger@mail.sub.uni-goettingen.de