



Spezifität und Multioptionalität

Kriteriologische Überlegungen zur Datenbankanalyse, -erstellung und -benutzung am Beispiel theologischer Fachdatenbanken

Christian Herrmann

Mit dem Übergang von gedruckten Bibliographien und Editionen zu CD-ROM-Datenbanken und schließlich zu elektronischen Publikationen im Internet stellt sich die Frage, inwieweit die gedruckten Ausgaben ersetzt werden können, welche zusätzlichen Informationen und Rechercheoptionen geboten werden und inwiefern die Beurteilungsmodi aus dem Bereich der qualitativ durchaus unterschiedenen Printformen auf die elektronischen Formen zu übertragen und zu modifizieren sind. Bei der Benutzung thematisch vergleichbarer Datenbanken erfährt man das Nebeneinander recht unterschiedlicher methodischer Konzeptionen der technischen Aufbereitung, aber auch inhaltlichen Gestaltung einerseits als Herausforderung und Anregung. Andererseits ist aber zu fragen, welche grundlegenden Problem- und Weichenstellungen vorhanden sind und ob die gebotenen Lösungsansätze adäquat gewählt wurden.

In drei Schritten soll ein Kriteriengerüst herausgearbeitet werden, das den Umgang mit Datenbanken in der bibliothekarischen Arbeit erleichtern könnte. Zur Exemplifizierung wurden die im Sondersammelgebiet Theologie / Religionswissenschaft der Universitätsbibliothek Tübingen verfügbaren Fachdatenbanken gewählt. Die *Theologie* ist insofern kein ganz ungeeignetes Beispiel, als hier historisch, sprachlich und - im Hinblick auf die Randgebiete - auch thematisch recht disparates Material zu verarbeiten und zudem das Nebeneinander primärer und sekundärer Quellen kennzeichnend ist, was für die Naturwissenschaften nur in einem geringen Maße zutrifft.

Im einzelnen handelt es sich um folgende Datenbanken: „Zeitschrifteninhaltsdienst Theologie“ (ZID)¹⁾, „ATLA Religion Database“ (RDB)²⁾, „South African Theological Bibliography“ (SATB)³⁾, „Francis“ (Francis)⁴⁾, „Patrologia Latina Database“ (PLD)⁵⁾, „Thomae Aquinatis opera omnia“ (Thomas)⁶⁾, „Archive of Celtic-Latin Literature“ (ACLL)⁷⁾, „CeteDoc“ (CETEDOC)⁸⁾, „Missio Nordica“ (MISSNORD)⁹⁾, „Religio“ (Religio)¹⁰⁾.

1. Authentizität als inhaltliche Intention

a) *Quantitative Authentizität*

Ziel einer Datenbank ist es, in einer möglichst authentischen, d. h. der u. U. komplexen Realität eines Sachverhaltes nachvollziehend gerecht werdenden Weise zu informieren. Da eine nur partielle, ausschnittshafte Rezeption der vorhandenen Informationsträger zu einer verengten, asymmetrischen Sichtweise und Erkenntnis führen kann, ist ein möglichst hohes Maß an *Vollständigkeit* in der Verfügbarmachung von Information anzustreben. Angesichts der unüberschaubaren Fülle an Informationsträgern und der gleichzeitigen Begrenzung technischer Speicherkapazitäten und nicht zuletzt der personellen und finan-

-
- 1) Hrsg. von der Universitätsbibliothek Tübingen / Theologische Abteilung, erscheint halbjährlich als CD-ROM.
 - 2) Hrsg. von der American Theological Library Association, erscheint jährlich.
 - 3) Hrsg. von der American Theological Library Association, erscheint jährlich.
 - 4) Hrsg. vom Institut de l'Information Scientifique et Technique (INIST, Frankreich), Aktualisierung ungefähr jährlich; vorausgesetzt wird hier wie auch sonst der Stand April 1999.
 - 5) Hrsg. durch den Verlag Chadwyck-Healey, basierend auf der Edition von Jacques-Paul Migne (erschienen 1844-1865 in Paris); vollständige Ausgabe auf 5 CD-ROMs 1995.
 - 6) Hrsg. von Roberto Busa bei EDITEL, Mailand; abgeschlossene vollständige CD-ROM-Ausgabe 1992, basierend auf der Printausgabe Busas „Opera omnia“ (7 Bde., 1980) und der Konkordanz „Index Thomisticus“ (57 Bde., 1974-1980).
 - 7) Hrsg. durch den Verlag Brepols (Turnhout) in Zusammenarbeit mit der Royal Irish Academy (Dublin), 1994; die Datenbank basiert auf diversen Printausgaben unterschiedlicher Provenienz und Qualität.
 - 8) Hrsg. von der Katholischen Universität Leuven, 3. Ausgabe 1996; die Datenbank gibt lateinische Kirchenvätertexte aus der Printausgabe „Corpus Christianorum, Series Latina“ sowie der zugehörigen „Continuatio Mediaevalis“ (erschienen bei Brepols, Turnhout) und einige weitere Texte aus anderen Ausgaben wieder.
 - 9) Angeboten über die Homepage der Staatsbibliothek in Århus:
<http://solwww4.statsbiblioteket.dk/www-bin/solscript/i=0,k=THVgT,l=1,c=base,a=7>
oder <http://www.misjonsnhs.no/MN.html>.
 - 10) Frei zugänglich im Internet, von einer Privatperson erstellt: <http://www.religio.de>.

ziellen Ressourcen in der Datenbankproduktion muss zwischen absoluter Vollständigkeit und relativer Relevanz vermittelt werden. So können Datenbanken erstellt werden, die thematisch eng eingegrenzt und damit nur für wenige Recherefälle relevant sind, in diesem engen Gebiet aber ein Höchstmaß an Vollständigkeit zu erreichen versuchen. Umgekehrt wird eine Ausweitung der thematischen Relevanz zumeist mit einer Modifizierung des Vollständigkeitspostulates zu dem Ziel benutzerorientierter Repräsentanz erkauf.

Die *quantitative Authentizität* lässt sich an der Wechselbeziehung der Größen Territorium, Sprache, Fach, Dokumenttyp, zeitliche Erstreckung festmachen. Das thematisch breiteste Spektrum bietet FRANCIS. Das bibliographisch referierte Material entstammt nicht nur den unterschiedlichsten Ländern und Sprachbereichen, sondern ist einer Vielzahl von Disziplinen (mit Ausnahme der Naturwissenschaften) zuzuordnen. Die *Multidisziplinarität* impliziert das Problem der fachlichen Einordnung bei Indexierung und Recherche. Eine Akzentuierung des französischsprachigen Bereiches kommt in einer Überrepräsentanz entsprechender Literatur¹¹⁾ und vor allem in der durchgehend französischen Sacherschließung zum Ausdruck, zu der nur sehr punktuell englisch-, deutsch- oder spanischsprachige Äquivalente treten. Die zeitliche Erstreckung ist mit dem irreversiblen Anfangspunkt 1984 recht gering. Die Auswahl der zu indexierenden Quellen scheint u. a. auch im Hinblick auf thematische Berührungspunkte zwischen den Disziplinen zu erfolgen, um so in Entsprechung zum inhaltlichen Datenbankprofil interdisziplinäre Studien zu erleichtern¹²⁾. Über den inhaltlichen Vergleich hinaus ermöglicht die Vielfalt der ausgewerteten Dokumenttypen (z. B. auch Forschungsberichte, Dissertationen, Rezensionen, Karten, Rechtsprechung, Gesetzgebung, graue Literatur) eine stilistisch-perspektivische Gegenüberstellung.

Das Beispiel der RDB, die sich auf die Bestände der protestantischen theologischen Bibliotheken der USA stützt und vor allem auf den angloamerikanischen Markt abzielt, zeigt, dass die Notwendigkeit der *Vermarktung* einer Datenbank zu akzentuellen Verschiebungen und Zuspitzungen des inhaltlichen Profils führen kann, die den Nutzwert für außerhalb des im Blick befindlichen Marktsegments Befindliche erheblich verringert. Zwar wird zu einem recht frühen Zeitpunkt (1952) mit der Indexierung begonnen, das zugrundeliegende Quellenreservoir aber nur sehr sukzessive ausgeweitet. Die Kurzrezensionen machen ein Drittel der erfassten Dokumente aus. Es ist ein deutlicher

11) So sind im Zeitsegment 1984-1990 187.042 französische, 168.550 englische, 55.229 deutsche, 23.627 spanische Titel u. s. w. enthalten.

12) Man vergleiche hierzu das Selektionsmenü zum Schlagwort-Element „connaissance“, das mit Schlagwörtern (und darüber vermittelt: Dokumentbeschreibungen) aus recht unterschiedlichen Fächern und methodischen Zugriffen verbunden ist.

*Anglozentrismus*¹³⁾ sowie ein Übergewicht *protestantischer* Publikationen festzustellen¹⁴⁾. Historische Entwicklungen wie die zeitweilige separate Indexierung methodistischer Publikationen werden nicht durch Integration revidiert, sondern hinterlassen störende und wegen partieller Doppellerschließung auch irritierende Erschließungsfragmente. Im ZID wird bei der weitgehenden Beschränkung auf den westeuropäischen Sprach-, nicht Territorialbereich weniger auf die Muttersprache möglicher Abonnenten, sondern auf das *quantitative und qualitative Gewicht* der entsprechenden theologischen Produktion geachtet. Sobald allerdings eine eigenständige, hochwertige und produktive theologische Forschung im osteuropäischen Raum etabliert ist, wird man Wege zu einer Ausweitung des Quellenreservoirs in diesen Bereich hinein finden müssen. Der ZID basiert auf den allerdings umfassenden Beständen einer einzelnen Bibliothek (UB Tübingen), die den Sondersammelgebietsauftrag der DFG für dieses Fach wahrnimmt. Da die Erschließung der Monographien über die elektronischen Katalogdatenbanken im Internet sowie monatlich in Neuerwerbungslisten zugänglich gemacht wird, kann sich diese Datenbank auf bibliographisch unselbständige Literatur beschränken¹⁵⁾. Eine Besonderheit stellt der flexible zeitliche Anfangspunkt dar, insofern sukzessive die gedruckten ZID-Hefte¹⁶⁾ von 1994 an rückwärtig mit allerdings reduzierter Sacherschließung in elektronische Form konvertiert und in die Datenbank integriert werden.

Ein *territorialer Zugriff* kann in unterschiedlichen Variationen betrieben werden. Die SATB informiert über die südafrikanische Produktion unabhängig von Sprache und Dokumenttyp, wobei allerdings ein Übergewicht des burisch-reformierten Spektrums und damit des in Afrikaans verfassten Schrifttums

13) Unter den vollständig indexierten Dokumenten sind 80 % in englischer Sprache verfasst (von 474.150 Titeln sind 419.290 englischsprachig); die nicht mit Sprachbezeichnungen versehenen Dokumentbeschreibungen (v. a. Rezensionen) betreffen zu 100 % englischsprachige Publikationen.

14) So erscheint Thomas von Aquin als Personenschlagwort bei 2.152 Titeln, Martin Luther dagegen bei 5.502 und Johannes Calvin bei 4.874 Treffern. Ähnliches ließe sich mit neueren Beispielen nachweisen (Karl Rahner: 735 Treffer; Karl Barth: 2.069) (Stand: 1999).

15) Aufsätze aus Zeitschriften, Festschriften, Kongressschriften (ab Erscheinungsjahr 1999), nicht aber aus thematischen Aufsatzbänden (wegen Kapazitätsmangel); eine Besonderheit stellen aufgrund vertraglicher Abmachungen die Dokumente der Deutschen Bischofskonferenz dar, die sowohl in bibliographisch selbständiger wie unselbständiger Form begegnen.

16) Seit 1975 wird monatlich über den Inhalt der jeweils neu hereingekommenen, wissenschaftlich relevanten Zeitschriftenhefte des SSG Theologie / Religionswissenschaft in Form eines Current-Contents-Dienstes informiert.

festzustellen ist. Eine Verknüpfung von territorialer und thematischer Eingrenzung begegnet bei MISSNORD: es wird alles geboten, was mit Mission bzw. Missionswissenschaft in den nordischen Ländern und von diesen ausgehend zusammenhängt (thematisch Einschlägiges aus Skandinavien, in skandinavischen Sprachen oder aus den Missionsländern, wenn ein Zusammenhang zu missionarischer Aktivität von Skandinaviern besteht). Das Territorium als Deskription der Subjektseite (Ort der Produktion), nicht der Objektseite (thematischer Gegenstand) kann aber auch abgesehen von einem Thema mit einer Sprache kombiniert werden. Dies trifft auf ACLL zu, wobei freilich das Lateinische in dieser Zeit (Antike, Frühmittelalter) als Schriftsprache üblich war¹⁷. Mit der sprachlichen Eingrenzung ist zugleich eine zeitliche Rahmenangabe verbunden, weil das Lateinische bereits zu dieser Zeit von muttersprachlichen Elementen durchsetzt und dann vollständig von anderen Sprachen auch als Schriftsprache abgelöst wird.

Bei den Volltextdatenbanken kann zwischen solchen unterschieden werden, die wie THOMAS alle Werke eines *einzelnen Autors* unter Zugrundelegung einer bestimmten Textedition wiedergeben, und solchen, die wie CETEDOC, PLD und ACLL Werke *unterschiedlicher Autoren* unter Verwendung diverser Texteditionen wiedergeben, wobei eine sprachliche und zeitliche Eingrenzung notwendig ist, um die Datenbank überschaubar zu halten¹⁸. Dabei werden auch Texte anderer Zeitepochen bzw. anderer Personen integriert, wenn diese unmittelbar historisch mit einem der Texte des festen Quellenkanons zusammenhängen (z. B. briefliche Korrespondenz; über einen späteren Kommentar oder eine Einleitung vermittelte Überlieferung). Das Nebeneinander unterschiedlichster Dokumenttypen beruht in seinem *Dass* auf dem Bemühen um vollständige Wiedergabe der überlieferten Texte, in seinem *So - d. h. in u. U.* lückenhafter Repräsentanz einzelner Textgattungen - freilich auf der tatsächlichen Produktion eines Autors.

Eine *Mischform* zwischen Referenz- und Volltextdatenbank stellt RELIGIO dar. Es werden sowohl von dem Redakteur selbst erstellte als auch von ihm zu thematischen Blöcken kompilierte fremde Volltexte unterschiedlichster Art (Dokumenttyp, stilistisches Niveau, Sprache, Ausführlichkeit, technische

17) Es wird hierbei mittels geographischer Codes zwischen verschiedenen Regionen des keltischen Kulturraumes differenziert.

18) Bei der PLD begegnet mit dem Zeitraum ca. 200 - 1216 n. Chr. ein sehr breit angelegtes Terrain, das nur durch die vergleichsweise geringe Literaturproduktion und lückenhafte Überlieferung aus dieser Zeit möglich ist. Die aufgrund des Textvolumens notwendige Aufteilung auf zwei physische CD-ROM-Scheiben macht zugleich wie im Fall von CETEDOC eine grobe zeitliche Eingrenzung (Antike / Vulgata und Mittelalter) möglich.

Strukturierung im Hinblick auf Querverweise mit Hilfe von Links) zusammengestellt. Der partiell bibliographische Charakter liegt darin, dass Listen mit eigenständig formulierten, thesenartig zusammenfassenden Links zu Volltexten verwendet und weiterführende Literaturhinweise gegeben werden. Die Vollständigkeit wird angesichts fehlender sonstiger Begrenzungen im Sinne der Exemplarität verstanden.

b) Formale Authentizität

Neben die quantitative tritt die *formale Authentizität*, die eine eindeutige Identifizierung der Autoren und Werke ermöglichen soll. Das Anliegen einer bis in die äußere Gestaltung und Schreibweise hinein die Vorlage nachbildenden Wiedergabe wird durchbrochen, um einen Text auch unabhängig von einer konkreten physischen Vorlageform wiedererkennbar und verifizierbar zu machen. Das Problem der Erstellung und Beobachtung von *Normierungsregeln* stellt sich weniger dort, wo wie bei Volltextdatenbanken eine Zeitepoche und Sprache so stark dominiert, dass alle abweichenden Formen dem übergewichtigen Schema subsumiert werden können¹⁹⁾. Aber auch in Volltextdatenbanken können Homonymenzusätze v. a. in Form regionaler Bezüge vergeben werden, um die zumeist unter dem Vornamen firmierenden Personen unterscheidbar zu machen²⁰⁾. In Referenzdatenbanken kann der Versuch, *Splittings* zwischen Einträgen mit ausgeschriebenen und abgekürzten bzw. unterschiedlich verbundenen doppelten Vornamen durch eine rigorose Reduktion der Vornamen zur Initialform zu vermeiden (FRANCIS, SATB), eine Verwechslung, weil Ineinanderordnung verschiedener Personen nicht verhindern. Abgesehen davon, dass es bei der Behandlung der Namenspräfixe trotzdem zu formalen Inkonsistenzen kommen kann (FRANCIS), ist es sicherlich komfortabler, über Trunkierungsverfahren differente Namensformen zusammenführen als erst gar nicht mit den ausgeschriebenen Vornamen recherchieren zu können. Zudem macht es einen Unterschied, ob die Splittings auf einem Fehlen konzeptioneller Einheitlichkeit beruhen und daher nicht eliminiert werden können (RDB) oder sukzessive im Verlauf der Datenbankeinstellung durch Na-

19) In den lateinischen Volltextdatenbanken werden alle Personennamen in der lateinischen Vorlageform angesetzt, die aus den gängigen kritischen Editionen bekannt sind. Personen späterer Zeiten (so in der PLD) werden in latinisierter Form wiedergegeben (z. B. „Schoenemannus, Fredericus G.“).

20) Dies ist insbesondere in ACLL zu beobachten.

mensabgleich gegenüber Normdatenbanken und Vervollständigung von Vornamen beseitigt werden (ZID)²¹⁾.

c) *Modale Authentizität*

In einem gewissen Gegensatz zu der von außen an die Vorlage herantretenden Normierung steht die Normierung durch die Vorlageform für die Gestaltung der Wiedergabe, also die *modale Authentizität*. Es geht letztlich um die Frage, inwieweit eine Volltextdatenbank im Hinblick auf die fortlaufende Lektüre die gedruckte Ausgabe bzw. eine Referenzdatenbank im Hinblick auf den Sachtitel das Inhaltsverzeichnis einer Zeitschrift oder die Haupttitelseite einer Monographie ersetzen kann. Dies betrifft erstens das Problem der *Sonderzeichen* bzw. Buchstaben aus nichtlateinischen Schriften. Hier muss eine an weitverbreiteten Regeln orientierte und daher eindeutige Transkription²²⁾ in der Datensatzanzeige, aber auch die Möglichkeit eines korrekten Datenexports geboten werden (so bei ZID)²³⁾. Elegant ist die Lösung, griechische Wörter in griechischer Form in einem Index zu suchen, wobei dieser Terminus nach der Übernahme als Suchbegriff automatisch transkribiert wird (PLD). ACLL gibt die ursprüngliche sprachliche Gestalt samt allen aufgrund keltischer Einflüsse zustande gekommenen Modifizierungen gegenüber dem klassischen Latein exakt wieder und kennzeichnet Bestandteile nichtlateinischer Sprachen durch Indikatoren. Zwar gewährleistet zweitens die Aufgliederung des Quelltextes in einzelne Sätze (*sententiae*) samt Referenzangaben (als „Memento“) (CETEDOC, ACLL) bzw. codierte Abschnitte (THOMAS) die *Zitierfähigkeit*²⁴⁾, aber eine *kontinuierliche Lektüre* - dritter Aspekt - , wozu auch und v. a. die gedruckten Editionen dienen, wird dadurch nicht gerade erleichtert. In ACLL

-
- 21) Der ZID orientiert sich bei Autoren an RAK (SWB Aut.), bei Personenschlagwörtern an der SWD, auch können bei einem Fehlen entsprechender Einträge in diesen Normdatenbanken Normierungen aufgrund der Häufigkeit in den vorliegenden Aufsätzen vorgenommen werden. Splittings bestehen nur zwischen Formen mit ausgeschriebenem und abgekürztem Vornamen, weil in der Anfangsphase der Datenbank die Vornamen abgekürzt wurden und sukzessive ergänzt werden müssen oder die Vornamen nicht ermittelbar sind. Dabei werden nur die beiden ersten Vornamen aufgenommen und nur der erste ausgeschrieben.
 - 22) Z. B. für griechische Buchstaben RAK, Anlage 5, Tab. 2, für hebräische Buchstaben DIN 31636.
 - 23) In der DOS-Version des ZID besteht die Möglichkeit eines Exports mit sämtlichen Sonderzeichen, in der DOS- und in der Windows-Variante die Option, eine korrekte Umsetzung von Sonderzeichen für Word Perfect zu initiieren.
 - 24) Diese ist auch nur eingeschränkt vorhanden, weil z. B. bei ACLL nur die Zeile des Satzbeginns, nicht aber die des Satzendes im Bezug auf die Printausgabe angegeben wird.

wird allerdings versucht, die redaktionellen Bemerkungen der Herausgeber der Printausgabe korrekt wiederzugeben und dabei durch eigene zu ergänzen²⁵. Unter Zuhilfenahme von Farb-, Schriftart- und Klammerungsvarianten erreicht PLD ein hohes Maß an Substitutionseffizienz im Bezug auf die gedruckte Vorlage²⁶, zumal zwischen zergliederter und fortlaufender Darstellungsform unterschieden werden kann.

Das Maß der modalen Authentizität hängt auch ab von der *technischen Strukturierung* der Datenbank. Windows-Versionen bieten durch das Nebeneinander nicht nur unterschiedlicher Initiierungsmodi für bestimmte Funktionen (Funktionstasten, Menü, Maus mit Icons), sondern v. a. durch die Wechselbeziehung mehrerer Fenster bessere Möglichkeiten zum Nachvollzug der potentiellen Benutzungsoptionen im Hinblick auf die Printversionen. Die Windows-Funktionalitäten können dabei mehr oder weniger stark ausdifferenziert sein: z. B. bietet PLD die Möglichkeit einer korrespondierenden Bewegung - via Leuchtbalken oder Markierungsfläche - in Ergebnislisten- bzw. Inhaltsverzeichnisfenster einerseits und Volltextfenster andererseits und in CETEDOC kann man mit Hilfe der Maus zwischen Anzeigearten und Bänden wechseln und es werden Kurzinformationen in Gestalt kleiner Fenster eingeblendet. Die Darstellung nichtlateinischer Schriften in ihrer ursprünglichen, nichttranskribierten Form ist nur in einer Windows-Version möglich (vgl. PLD). Eine die Windows-spezifischen Link-Techniken in DOS-basierter Form vorwegnehmende Quasi-Hypertextfunktion (Aufruf einer mit dem Cursor markierten verbalen Umschreibung eines anderen Textes mit der Enter-Taste) begegnet bei THOMAS²⁷. Der Unterschied zwischen DOS- und Windows-Versionen wirkt

25) Bzgl. Hrsg. der Printausgabe: Ergänzungen (Winkelklammern), Streichungen (eckige Klammern), Marginalien / Glossen (Schweifklammern); bzgl. Datenbank-Hrsg.: Korrekturen (Kapitalchen), Kommentare (ebenfalls Schweifklammern); ähnlich: direkte Rede in Vorlage (Kommata), biblische Zitate in Vorlage (Doppelkommata in invertierter Form).

26) Kursiv-, Fett- und griechischsprachiger Druck wird in der Datenbank analog dargestellt. Lacunae (Lücken) werden gekennzeichnet („[...]“). Bezugnahmen auf biblische, patristische oder antik-klassische Werke, wie von Migne verzeichnet, erscheinen violett. Mignes Spaltenzahlen werden in blau und mit eckigen Klammern wiedergegeben, die Originalseitenzahlen des von Migne verwendeten Quelldokuments in Fettdruck.

27) Dies geht über den insbesondere von dem Aufruf der Menüfunktionen her bekannten Weg einer Kombination von Shift und farbig unterlegtem Buchstaben hinaus. In der DOS-Version des ZID wird außer einer automatischen Ausführung von Verweisungen bis hin zum Sprung in einen anderen Index die Möglichkeit geboten, über ein mit Hilfe eines „x“ aufzurufendes Fenster einen in einem Schlagwortstammsatz genannten verwandten Begriff aufzusuchen und zur ersten Eintragung dieses Schlagwortes im Schlagwortindex zu gelangen.

sich dort um so schmerzlicher aus, wo wie bei ACLL und CETEDOC bzw. SATB und RDB eine Kontinuität in Grundfunktionalitäten (gleiche Software-Firma!) vorhanden ist, ein Wechsel zwischen beiden Versionen mit ihren Spezifika und Funktionalitätsgewinnen aber nicht wie im ZID²⁸⁾ möglich ist. Eine Sondergestalt einer modalen Authentizität ist in RELIGIO festzustellen: hier gilt es weniger, eine gedruckte Ausgabe bis in die Aufmachung hinein wiederzugeben, sondern der Benutzer kann selbst in die Gestaltung der interaktiv verstandenen Datenbank per integrierter E-Mail-Funktion eingreifen und eigene Quellen einbringen, was freilich zu einer äußeren Inkonsistenz führt.

d) *Historische Authentizität*

Bei Volltextdatenbanken muss das Problem der *historischen Authentizität* beachtet werden, d. h. die Frage, inwieweit Erwägungen über die Richtigkeit der traditionellen Zuordnung von Autor und Werk in der Datenbank Berücksichtigung finden muss. Um der Asymmetrie zugunsten der modalen Authentizität zu entsprechen, unterscheidet die PLD zwischen Einschätzungen im Hinblick auf die Verfasserschaft einer Schrift, die von den Herausgebern der zugrundeliegenden Printausgabe selbst stammen²⁹⁾, und solchen, die von späteren Kommentatoren geäußert wurden³⁰⁾. CETEDOC setzt in Zweifelsfällen einen entsprechenden Vermerk („(dubium)“) hinter den Namen und ordnet anonyme Schriften nach Zeit, Adressat, Bezugspunkten, Textgattung oder direkt unter dem Sachtitel.

28) Die DOS-Version des ZID bietet anders als die Windows-Version die Option eines Exports mit sämtlichen Sonderzeichen, die Windows-Version dagegen ermöglicht eine Recherche im Expertenmodus und zahlreiche Textverarbeitungsfunktionen an der Datensatzanzeige. Letztere müssen aber als solche nicht zwingend an eine Windows-Version gebunden sein, wie die allerdings defizitäre Möglichkeit eines Copy- und Paste-Verfahrens in der SATB zeigt.

29) Wenn bereits bei Migne die Verfasserschaft nicht geklärt oder unbekannt ist, wird die Formulierung „Auctor incertus“ verwendet oder ein vermuteter Autorenname in Klammern und mit Fragezeichen versehen dahinter gesetzt (z. B. „Auctor incertus (Patricius Hibernae?)“). „Auctores varii“ deutet auf die Beteiligung mehrerer Autoren hin, „Editores“ auf die Kompilation eines Textes durch Migne oder seine Verleger.

30) In diesem Fall werden die Dokumente mit einem Code (in den Suchergebnissen und über dem Volltext) versehen, der aus einem oder mehreren von drei möglichen Buchstaben als Abkürzungen für neuere Kommentare zur Migne-Ausgabe stehen. Ein Sternchen neben dem Code deutet darauf hin, daß in diesen Hilfswerken nicht nur auf die jeweilige Schrift Bezug genommen, sondern die Autorenangabe Mignes in Frage gestellt wird.

2. Deduktive und induktive Hierarchie als Recherchemodus

Der Zugriff auf einen Text kann einerseits *sequentiell* erfolgen, indem man einen Volltext oder ein Inhaltsverzeichnis einer Zeitschrift kontinuierlich von Anfang an liest, bis man auf den gesuchten Sachverhalt stößt. Dieses Verfahren hat den Vorteil einer größtmöglichen Beziehbarkeit auf den Kontext, aber den Nachteil eines sehr hohen Zeitaufwandes. Andererseits kann man *punktuell* auf bestimmte Stellen zugreifen, in denen ein Suchbegriff vorkommt. Eine gedruckte Ausgabe bietet dabei einen eindimensionalen, d. h. auf Kombinationen verzichtenden, und sukzessiven Zugriff: die Einzeltexte müssen aufgrund weiterer mangelnder Eingrenzungsmöglichkeiten nacheinander auf Relevanz hin befragt werden. *Eine Datenbank erlaubt eine mehrdimensionale und simultane Punktualisierung*: durch zahlreiche Kombinationsmöglichkeiten einzelner Suchbegriffe können gleichzeitig sämtliche Teile eines Textes nicht nur auf ein Vorkommen, sondern auf ein *relevantes* Vorkommen eines Suchbegriffes abgeprüft werden. Ein sequentielles Vorgehen im Sinne einer fortlaufenden Lektüre impliziert dabei im Hinblick auf den einzelnen Suchbegriff die größtmögliche Punktualität, weil dieser Begriff nur an dieser Stelle in dem Kontext der konkreten anderen Wörter, der Satzkonstruktion, der intendierten Sinnggebung und Position innerhalb der Gliederung steht. Umgekehrt kommt ein punktuellles Vorgehen nicht ohne sequenzialisierende Verfahrensschritte aus, weil der Einzelbegriff mit abstrakter Lokalisierung in Form dezimaler Gliederungspunkte in der Regel nicht zitierbar und ohne Rekurs auf einen weiteren Kontext nicht voll verstehbar ist.

Diese zunächst aus der Beobachtung der Benutzungsweise einer gedruckten Edition, aber auch eines in physischer Form vorliegenden Zeitschriftenjahrganges gewonnene Wechselbeziehung der Termini „sequentiell“ und „punktuell“ wird zumeist im Bezug auf Datenbanken durch die Verhältnisbestimmung der Begriffe „*precision*“ und „*recall*“ modifiziert³¹⁾. Hierbei geht es allerdings eher um eine Beschreibung der Qualität der Suchergebnisse als um eine Definition des Rechercheverfahrens. Ein hohes Maß an „*precision*“, d. h. enger semantischer Füllung, impliziert ein geringes Maß an „*recall*“, d. h. Quantität des Vorkommens. Es kann dabei sein, dass semantisch etwas weiter gefasste, dann auch zahlreichere Begriffe durchaus noch in den Bereich des Relevanten fallen würden, eine Verringerung der „*precision*“ also wün-

31) Z. B. Klaus Lepsky/Jörg Siepmann/Andrea Zimmermann: „Automatische Indexierung für Online-Kataloge: Ergebnisse eines Retrievaltests“. - In: ZfBB 43 (1996), S. 47-56, hier: S. 54f.; Elaine Svenonius: „Präkoordination - ja oder nein?“. - In: ZfBB 41 (1994), S. 279-296, hier: S. 281.284.290f.; Kommission des Deutschen Bibliotheksinstituts für Erschließung und Katalogmanagement, Expertengruppe Online-Kataloge: Sacherschließung in Online-Katalogen. - Berlin, 1994, S. 44

schenswert sein könnte. Durch das *Extraktionsverfahren* in Gestalt der Stichwortsuche in Volltexten oder Sachtiteln wird die „precision“ zusätzlich durch die Verwendung bestimmter grammatikalischer Formen oder Kompositabilidungen - bei häufig gleichbleibender Semantik - erhöht. In einem Volltext ist ein einzelnes *Wort vielfach definiert* durch einen bestimmten Gliederungskontext, aber auch durch die unmittelbare Umgebung anderer Wörter und die Integration in eine Satzkonstruktion mit ihren Folgen für die grammatikalische Flexionsform, wobei der formale Rahmen inhaltliche Aussagen implizieren kann. Da zumal bei der Suche in einem Volltext ein großes Textvolumen als Suchraum verwendet wird, muss eine Eingrenzung der Suchergebnisse auf die konkrete Form und spezifische Bedeutung und somit der *aposteriorische* Nachvollzug der *apriorischen* Begrenztheit des Begriffes möglich sein. Andererseits müssen Verfahren zur Verfügung stehen, die zumindest die für die Semantik unerhebliche *formale Variationenbreite* relativieren und Abstufungen in dem Maß der semantischen Definition zulassen.

Um sowohl eine Verknüpfung zu den Printausgaben zu erreichen als auch die intentionale Zielsetzung des Zugriffs im Sinne der Qualität des Suchergebnisses zum Ausdruck zu bringen, bietet sich *zur Beschreibung und Beurteilung der Rechercheoptionen von Datenbanken eine spezifische terminologische Differenzierung* an. Wie noch zu zeigen sein wird, gehen Datenbanken stets in hierarchischer Weise vor. Um sowohl einen sequentiellen wie punktuellen Zugriff zu gewährleisten als auch „precision“ und „recall“ zu erhöhen oder zu verringern, betreiben Datenbanken eine *deduktive* und *induktive* Hierarchisierung.

So wie man bei einem mehrbändigen gedruckten Werk zunächst einen Band auswählt, dann das Titelblatt, das Inhaltsverzeichnis mit Teilüberschriften studiert, schließlich eine bestimmte Seite oder Spalte aufschlägt und dort zu lesen beginnt und nach einem Begriff bzw. begrifflich explizierbaren Sachverhalt sucht, so wird in einer Volltextdatenbank idealerweise ein die formale Gliederung des Textes nachahmendes *Subordinationsverfahren* angeboten. Es geht um eine *Deduktion*, d. h. dass eine zunehmende Einengung des äußeren Rahmens eines Begriffes betrieben wird. Dabei lässt man sich von der Erkenntnis leiten, dass einer Formulierung in einer Titelangabe ein höheres Gewicht zukommt als einer solchen in einem ausführenden Text und dass *thetisch vorgetragene Sätze anders zu interpretieren sind als in einer polemischen Auseinandersetzung vorgebrachte*. Ein solches Vorgehen wird in der PLD durch eine Unterteilung des Textes in *Gliederungselemente* und eine Kombinationsmöglichkeit von verbalen Suchbegriffen und abstrakten Gliederungselementen angeboten. Die Recherche mit Hilfe sogenannter Kontext-

operatoren³²⁾ („in“, „nicht in“, „enth“) ist die sprachliche Explikation der hierarchischen Struktur des Datenbankaufbaus. Dabei lässt sich jedes Element in eines oder mehrere ihm zu- und untergeordnete Elemente unterteilen, die sich in ihrer Benennung an der SGML-Codierung orientieren³³⁾. Der Ausgangspunkt kann entweder im Suchwort liegen („in“) oder im Element („enth“), wobei letzteres Verfahren zu einer geringeren Trefferzahl führt, weil ein Suchwort in einem Element mehrmals vorkommen kann. Wenn nicht nach bestimmten Begriffen gesucht wird, sondern eine an einer bestimmten Stelle in der Gliederung einsetzende fortlaufende Lektüre angestrebt wird, kann man sich in PLD mit Hilfe des Inhaltsverzeichnisses, das Weiterverzweigungen durch ein Pluszeichen andeutet, an die gewünschte Stelle durch das Click-Verfahren hinbewegen³⁴⁾. Bei THOMAS wird zwar ebenfalls eine Kombination von Gliederungseinheit und Suchbegriff angestrebt, aber die Perfektionierung der modalen Authentizität zu Lasten einer abstrahierenden Normierung macht den tatsächlichen Gebrauch dieser Option fast unmöglich. Man muss die konkreten Bezeichnungen der jeweiligen Gliederungseinheiten in dem betreffenden Einzelfall kennen³⁵⁾, um eine Eingrenzung mit Hilfe einer Zuordnung zur Gliederung betreiben zu können. Eine normierende, d. h. generische Bezeichnung³⁶⁾ wäre hier hilfreicher. In CETEDOC ist nur eine Unterscheidung von Titel- und Volltextstichwort möglich. Eine Subordination in dem Sinne, dass ein verbaler Suchbegriff mit einer formalen Größe als Suchraum verbunden wird, liegt auch dann vor, wenn in allen Datenbanken entweder ein extrahiertes Volltext- oder Titelstichwort bzw. ein intellektuell vergebenes Schlagwort mit der Angabe eines konkreten Autors, eines Erscheinungsjahres, Verleges, Ortes, einer Sprache u. s. w. als Suchanfrage verknüpft werden kann. CETEDOC

32) Die PLD unterscheidet zwischen „Kontextoperatoren“, die sich auf die Gliederung beziehen, und „Umgebungsoperatoren“, die die Wörter in der unmittelbaren Umgebung eines Suchbegriffes in einem Satz im Blick haben.

33) SGML = Standard Generalised Mark-Up-Language; ein <volume> besteht z. B. aus <volfront>, <volbody> und evtl. <volback>; ein <volbody> enthält eine oder mehrere <group>s und jedes <volback> eine oder mehrere <doc>s und / oder <group>s; jedes <doc> ist in Teile (<div?>s) und Abschnitte (<p>) aufgeteilt; man kann die Suche auf den Anmerkungsteil beschränken, wenn man das Kennzeichen <note> verwendet.

34) Zu den Recherchemöglichkeiten in der PLD: Tony O'Rourke: „Migne auf CD-ROM: Erstellung der Patrologia Latina Database (Stand: Mai 1994)“. In: Mitteilungsblatt der Arbeitsgemeinschaft Katholisch-Theologischer Bibliotheken 41 (1994), S. 99-104, bes. S. 102f.

35) Z. B. Summa theologiae I-II, q. 31, a. 2, ad 3; in anderen quaestiones treten möglicherweise wieder völlig andere Untereinheiten auf.

36) Also z. B. Hierarchiestufe 1, 2, 3 u.s.w.

und ACLL bringen dies sachgemäß mit der Unterscheidung von „formae“, d. h. konkreten Einzelbegriffen aus dem Volltext, und „Filtern“, also äußeren Eingrenzungen wie Autor, Werk, Zeitraum, zum Ausdruck³⁷⁾. Die Bestimmung solcher äußeren, nicht auf die Gliederung bezogenen Definitionsgrößen als Suchraum erfolgt in der PLD als Festsetzung eines „Attributes“: aus einer Zahl koordinierter Größen, z. B. allen Bänden einer Werkausgabe oder allen in der Datenbank vorkommenden Autoren, wird ein Konkretum selektiert und benannt³⁸⁾. Diese erste Variante der deduktiven Hierarchisierung lässt sich durch die *Inbeziehungsetzung von nicht auf der gleichen Stufe Stehendem* kennzeichnen, was in der PLD auch die Unmöglichkeit einer gleichzeitigen Suche nach Suchbegriffen und Gliederungselementen mit sich bringt³⁹⁾.

Ein deduktives, weil einschränkendes, präzisierendes Vorgehen stellt auch die Verwendung von Umgebungsoperatoren⁴⁰⁾ und Booleschen Operatoren dar. Hierbei wird ein Suchbegriff durch seine Positionierung im Rahmen bestimmter Formulierungen bzw. durch syntaktische Verknüpfungen näher definiert. *Umgebungsoperatoren* können in Referenzdatenbanken angeboten werden, wobei hier die Bezugnahme auf Kategorien als Rahmen des gemeinsamen Vorkommens zweier Suchbegriffe („same“) und die v. a. im Hinblick auf Autorennamen wichtige assoziative Suche („like“) als Spezifikum auftreten dürfte (so bei RDB⁴¹⁾). Da allerdings nur der Sachtitel die Textgrundlage bildet, während die gelegentlich aufgenommenen Inhaltsverzeichnisse von Festschriften nicht recherchierbar sind und keine Summarien erstellt werden, bleibt der

37) In Referenzdatenbanken entspricht dies dem Verhältnis von Schlagwort-, z. T. auch Basic-Index-Feld und den übrigen Suchfeldern einer Suchmaske.

38) Z. B. „<volume n=“120“>“ oder „<author=“Schoenemannus, Fredericus G.“>“

39) Nicht möglich ist also eine Anfrage „haeres* oder <note> in <p>“; dagegen kann sehr wohl nur nach Elementen, nur nach verbalen Suchbegriffen oder nach Suchbegriffen innerhalb bestimmter Elemente bzw. nach Elementen, die bestimmte Suchbegriffe beinhalten, gesucht werden.

40) Ich verwende diesen Begriff im Sinne der PLD. Hiermit ist das Verhältnis der Wörter eines Satzes gemeint, nicht der Bezug auf die Gliederung des Textes. In der Regel spricht man mangels Suchmöglichkeiten über Gliederungselemente hier von „Kontextoperatoren“.

41) Die assoziative Suche ermöglicht z. B. bei Eingabe des Namens „Mayer“ auch Tref-fer unter „Meier“ und „Meyer“. RDB unterscheidet zwischen „same“ bzgl. Suchfeld und „sames“ bzgl. Satz. Typisch ist auch der Operator „none“ (Suche nach allen Dokumenten, in denen dieses Suchfeld leer ist) und „all“ (Suche nach allen Dokumenten, die Daten in diesem Suchfeld enthalten). Die Aussage der maximalen Entfernung zwischen zwei Suchbegriffen („near x“ bzw. „nears x“ [hier ist Reihenfolge relevant]) gehört dagegen zum Standardrepertoire der Umgebungsoperatoren.

Nutzen begrenzt⁴²⁾. In Volltextdatenbanken ist die Arbeit mit Umgebungsoperatoren, gerade auch durch die Möglichkeit, die Reihenfolge des Vorkommens zweier Begriffe innerhalb eines Entfernungsrahmens vorzugeben, aufgrund des großen Textvolumens und des Fehlens normierender, abstrahierender Schlagwörter hilfreich⁴³⁾.

In allen etwas komplexeren und umfangreicheren Datenbanken (Ausnahme hier: Religio) kann mit den *logischen (Booleschen) Operatoren* „und“, „oder“, „nicht“ eine Verknüpfung zwischen zwei Suchbegriffen auf derselben hierarchischen Ebene betrieben werden, wobei diese Modifizierung der Anfrage einschränkend oder ausweitend wirken kann. Die verbalisierte Form ist in der Regel in der sogenannten Expertensuche zu verwenden, die abstrakte Form in der Gestalt mathematischer Symbole, die entweder wie in der DOS-Version des ZID selbst in einen Index einzugeben oder in Windows-basierten Datenbanken aus zu einzelnen Suchfeldern gehörenden Menüs auszuwählen und per Mausclick zu initiieren sind⁴⁴⁾. Zumeist ist innerhalb der Suchfelder und zwischen den Suchfeldern ein bestimmter Verknüpfungsmodus *voreingestellt*, der verändert werden kann, während in davon unterschiedenen Feldern für die ungefilterte Suche nur eine verbale Eingabeform möglich ist⁴⁵⁾. Bei der im Hinblick auf das Rechercheverfahren rein verbal-produktiven Suche nach „formae“ kommt es zu einer allerdings reduzierten *Integration des Expertenmodus' in die Suchmaske*⁴⁶⁾. Es kann durch eine geschickte Klammerungstechnik eine komplexe, hierarchische Bearbeitungsreihenfolge vorgegeben werden. Allerdings kann man sich in dieser quasi *instrumentellen Hierarchie*

42) Immerhin ist unter den analysierten Referenzdatenbanken RDB die einzige, die eine Recherche mit Hilfe von Umgebungsoperatoren anbietet. Für FRANCIS werden zwar Summarien erstellt, aber deren Inhalt ist nur über den Basic Index in Form von Einzelbegriffen oder logischen Begriffskombinationen suchfähig.

43) So bei PLD, THOMAS, CETEDOC und ACLL.

44) So in Win-ZID (im Ergebnismengenfenster), in PLD verbale Eingabe und Auswahl als Zeichen möglich, in CETEDOC Zeichen, in ACLL im Formae-Feld nur verbale Eingabe.

45) Die voreingestellten Werte sind dabei durchaus verschieden: in THOMAS innerhalb der Suchfelder „and“, zwischen den Suchfeldern „or“; in CETEDOC innerhalb der Filter unveränderbar „und“ vorgegeben; in ACLL zwischen den Suchfeldern „und“ voreingestellt, ein „oder“ oder „und nicht“ auch nur bei Angabe des Suchraums für die Formae-Suche, nicht für die Lektüre sinnvoll; in SATB in den Suchfeldern „or“ vorgegeben, zwischen ihnen „and“; in MISSNORD ist der Operator „not“ nicht vorgesehen; in Win-ZID zwischen den Feldern „und“ voreingestellt, aber veränderbar.

46) So bei CETEDOC, ACLL; in THOMAS können nicht nur Boolesche, sondern auch Umgebungsoperatoren unmittelbar in die Suchfelder eingegeben werden, eine Klammerungstechnik ist aber nicht vorgesehen.

nur auf die „formae“, also Bestandteile der letzten Hierarchiestufe beziehen, aber weder wie bei einer ausgefeilten Expertensuche die einzelnen Suchfelder in Form verbaler Kürzel⁴⁷⁾ noch mögliche Gliederungselemente mittels entsprechender Operatoren⁴⁸⁾ einbeziehen. Auf der anderen Seite bietet die unterstützte Suche nicht nur den Vorteil, auf die Eruiierung der Kürzel für die Suchfelder und möglicher Operatoren zunächst verzichten zu können, sondern auch weitere nützliche Optionen⁴⁹⁾.

Ein hohes Maß an vorgängiger, bereits vollzogener und nicht erst durch den Rechercheur sukzessiv zu betreibender deduktiver Hierarchisierung liegt immer dann vor, wenn sich Referenzdatenbanken in ihrem Sacherschließungsmodus an der angloamerikanischen Methode der massiven *Präkombination* orientieren⁵⁰⁾. Hierbei werden in ihrem Bedeutungsgehalt eher weite Grundschlagwörter (main heading) nicht durch syntaktische Kombination mit separaten anderen Schlagwörtern (Präkoordination), sondern durch die Integration eingrenzender, aspekthafter Elemente wie Personen, Fachdisziplinen oder Zeitabschnitte als internen Konstitutiva (qualifier) präzisiert. Der Aussagegehalt ist dann so stark eingeeengt, dass diese präkombinierten Ausdrücke ohne eine innere Beziehung gleichordnend nebeneinander gestellt werden können und zumeist auf weitere Kombinationsmöglichkeiten etwa von verbaler und klassifikatorischer Erschließung verzichtet wird⁵¹⁾. Die Schwierigkeit, das sach-

47) So bei Win-ZID, FRANCIS (hier zugleich größere Zahl von Suchfeldern als bei der unterstützten Suche mit Hilfe einer Suchmaske; zur Angabe eines Zeitraumes können außerdem eckige Klammern verwendet werden).

48) So bei PLD

49) Z. B. die sog. „Cross reference search“, d. h. die Möglichkeit, aus einer Vollanzeige einzelne Bestandteile zu kopieren und in das entsprechende Suchfeld zu übernehmen (bei SATB allerdings nur mit ganzen Zeilen möglich, was allenfalls bei den untereinander aufgeführten Schlagwörtern sinnvoll sein dürfte); bei THOMAS die Funktion „calculate“, die die zu erwartende Trefferzahl vorweg aufführt, ohne dass die eigentliche Suche samt Trefferanzeige eingeleitet würde; technisch damit verwandt ist die aposteriorische Schnittmengenbildung bei ACLL: es wird zunächst für jedes von mehreren ausgefüllten Suchfeldern eine eigene Suche samt Anzeige der Trefferzahl durchgeführt und erst danach eine Kombination der so gebildeten Teilergebnismengen betrieben.

50) So in FRANCIS (hier v. a. mit Präzisierungen in Klammern, nicht mit Binde- oder Schrägstrich dahinter), RDB, SATB, MISSNORD (hier ist als Konsequenz des inhaltlichen Profils für die Sacherschließung die Möglichkeit einer Suche mit Hilfe von „geographical keywords“ zu nennen (als Präkombination von Kontinent und Land).

51) Unter den Datenbanken dieses Typs bietet nur FRANCIS aufgrund ihres multidisziplinären Charakters eine auf die Wissenschaftsfächer und Teildisziplinen bezogene Grobklassifikation an, die allerdings nicht verbalisiert und auch nirgendwo in decodierter Form aufgelistet wird.

liche Deskriptionselement in diesem hohen Präzisionsgrad für eine effiziente Suche bereits kennen zu müssen, wird zumeist durch ein hohes Maß an formaler und methodischer Inkonsistenz der Sacherschließung verstärkt⁵²⁾. Im ZID wird in Anlehnung an die SWD - bei einigen um der größeren Konsistenz geschehenden Abweichungen v. a. bei Textschlagwörtern - hingegen auf *Normierung* (Schlagwortstammsätze mit Verweisungen und Erläuterungen in zweifelhaften Fällen) Wert gelegt. Der apriorischen Präkoordination im Sinne einer *syntaktischen Indexierung*, die eine eigenständige Verwendung von eher weiten Aspektschlagwörtern durch ihre Einbindung in einen semantischen Kontext sinnvoll macht, entspricht erstens die aposteriorische *Postkoordination in der Suche*⁵³⁾. Zweitens gibt das durch die Kettendarstellung und -permutation im Index ermöglichte assoziative *Browsing* in den Schlagwortketten eine häufig bereits ausreichende Information im Hinblick auf die Relevanz eines so beschriebenen Dokuments⁵⁴⁾, ohne dass die Vollform der Datensatzanzeige aufgerufen werden müsste.

Der Nutzen einer Datenbank ist daran zu messen, inwieweit den Nachteilen des deduktiven Suchverfahrens entgegengewirkt wird. Als eine *Induktion in der Deduktion* kann ein Vorgehen bezeichnet werden, das zwar grundsätzlich

52) In allen genannten Datenbanken wird nicht zwischen Verweisungen und Vorzugsbenennungen unterschieden. Die Variationenbreite möglicher Einstiege spiegelt sich partiell in dem Nebeneinander mehrerer in ihrer Bedeutung weitgehend identischer Schlagwörter wieder, die weder subordinativ (Verweisung auf Ansetzungsform) noch assoziativ (Querverweis auf verwandte Begriffe) aufeinander bezogen werden. In der RDB sind biblische Textstellen sowohl im Suchfeld „scripture reference“ (standardisierte Abkürzungen, die man aber kennen muss) als auch als Schlagwort (subject heading), dort aber nur als Unter-Schlagwort zu „Bible (OT)“ bzw. „Bible (NT)“ suchbar. Predigten werden in der RDB nicht unter Hinweis auf den ihnen zugrundeliegenden Text erschlossen, was eine Recherche nach bestimmten „sermons“ nahezu unmöglich macht. Abstracts werden nicht konsequent für alle Dokumente und uneinheitlich in ihrer Ausführlichkeit (oft nur Satzform der ohnehin in sich schon mehrgliedrigen Schlagwörter) angeboten (FRANCIS, RDB, SATB). Defizite wie die tatsächlich fehlende sachliche Erschließung in zahlreichen Datensätzen (erkennbar an Sortieroption, nichtindexierte Dokumente mit anzuzeigen: RDB, SATB) haben allerdings häufig ihren Grund in der historischen Genese der Datenbank. Dies könnte auch für die Differenz in der Erschließung rezenter Dokumente und der Retrokonversion der gedruckten Hefte (nur Personen- und Textschlagwörter, die im Sachtitel enthalten sind; ohne Autopsie) im ZID gesagt werden.

53) Entweder mit Hilfe der logischen Operatoren bei Zerschlagung der Kettendarstellung mit Hilfe der sog. variablen Trunkierung oder durch das Hineinschreiben einer Kette in das Suchfenster für Index 4 (DOS-Version) bzw. das Fenster für die Expertensuche (Windows-Version) (z. B. „Abendmahl ; Neues Testament“).

54) Zur Begrifflichkeit von Präkoordination, Postkoordination, Präkombination vgl. DIN 31623, Teil 2: Indexierung zur inhaltlichen Erschließung von Dokumenten.

im Bereich der deduktiven, also einengenden Suche seinen Ort hat, hierbei aber ausweitend wirkt. So ist in der RDB und SATB eine „*Thesaurus search*“ möglich, in der - dem punktuellen Zugriff auf ein bestimmtes Dokument vorgelagert - zunächst dem eingegebenen Suchbegriff verwandte Schlagwörter (z. B. sämtliche Präkombinationen zu einem Grundschlagwort) aufgefunden werden, um dann erst in einem zweiten Schritt („*Topical search*“) zu den mit den gefundenen Schlagwörtern verbundenen Dokumentbeschreibungen zu gelangen⁵⁵⁾. In FRANCIS wird, wenn nur mit Hilfe eines Schlagwortes gesucht und dieses in einigen Variationen, d. h. Präkombinationen verwendet wird, ein *Selektionsmenü* quasi als Zwischenstation geboten, in dem die möglichen Formen koordiniert aufgelistet sind, um so einen gezielten, deduktiven Zugriff zu erleichtern⁵⁶⁾. Das Problem einer auch bei Zuhilfenahme der Indizes bleibenden Unkenntnis der präkombinierten Schlagwortformen, aber auch komplizierten Körperschaftsbezeichnungen versuchen viele Datenbanken durch eine Differenzierung von „*word*“-*Index* (Einzelbestandteil) und „*entry*“-*Index* (Vollform) zu beheben⁵⁷⁾. Im ZID werden die Ziele dieser Verfahrensweisen durch einen auch in der Windows-Version bleibenden prioritären Bezug auf die Indizes erreicht, da hier über eine konsequente Verweisungstechnik potentielle Variationen des Sucheinstiegs aufgefangen, durch die variable Trunkierung und den Kreuzindex Einzelschlagwörter bzw. Einzelbestandteile aus formalen und sachlichen Deskriptionselementen sowie die subordinierten Teile von Ansetzungsketten (z. B. Werkittel bei Personenschlagwörtern) in der Gestalt von Verweisungsformen suchbar gemacht werden⁵⁸⁾.

- 55) Die gleiche Funktion erfüllen im ZID die in die Schlagwortstammsätze integrierten Siehe-auch-Verweisungen, die ebenso wie die in den Dokumentbeschreibungen verwendeten Schlagwörter in der Windows-Version als Link (in Allegro-Sprache „Flip“) anklickbar sind und zu dem ersten entsprechenden Indexeintrag, d. h. inhaltlich verwandten Dokumenten, hinführen.
- 56) Vergleichbar ist es, wenn bei einer reinen Titelstichwortsuche in CETEDOC zunächst sämtliche Schriften aufgelistet werden, die diese Titelbestandteile enthalten. Über eine Deselektion können die für die fortlaufende Lektüre als nicht relevant empfundenen Schriften aus der Treffermenge entfernt werden.
- 57) Bei Volltextdatenbanken wie CETEDOC / ACLL ist dies v. a. bei Sachtiteln nützlich, deren Elemente allerdings auch als Suchbegriffe unmittelbar in das entsprechende Suchfeld eingetragen werden können. Ein Problem besteht, wenn wie in FRANCIS bei Übernahme mehrerer Begriffe aus dem Index eine „or“-Verknüpfung vorgegeben ist und das betreffende Suchfeld zu klein ist, um eine Kombination mehrerer aus dem Index übernommener Einzelelemente zu ermöglichen (bei RDB und SATB ist diese Schwierigkeit nicht in diesem Maße gegeben).
- 58) Im ZID ist allerdings aus vorgegebenen technischen Gründen eine Stammsatzverknüpfung, d. h. eine unmittelbare Suche nach Dokumentbeschreibungen durch Eingabe auch der Verweisungsformen nicht möglich. Man gelangt von Verweisungsformen stets zunächst zu der betreffenden Ansetzungsform und von dort zu den damit verknüpften Datensätzen.

Eine *induktive Hierarchisierung* liegt erstens dann vor, wenn von einem Einzeltreffer aus der Kontext und damit der Bezugsrahmen erweitert wird. Das geschieht in der PLD über die „Gehe-zu“-Funktion, die den kleinen in der Ergebnisliste angezeigten unmittelbaren Kontext eines gesuchten Begriffs zu einem die fortlaufende Lektüre größerer Abschnitte ermöglichenden Kontext ausweitet. In CETEDOC geschieht dies durch die Unterscheidung von Einzelsatzanzeige (*sententiae*) und fortlaufendem Volltext (*textus*). In Win-ZID kann von der Vollanzeige eines einzelnen Datensatzes aus zu weiteren Einträgen unter dem Autor oder unter jedem der verwendeten Schlagwörter „geflippt“ werden.

Zweitens kommt es zu einer Ausweitung des Bezugsrahmens, wenn nicht auf den größeren Kontext der Verwendung desselben Suchbegriffs (*modal-quantitativ*), sondern auf die Variationenbreite des Suchbegriffes als solchen abgehoben wird (*modal-qualitativ*). Dies kann erstens mit Hilfe der Trunkierung als Ersatz mehrerer Buchstaben bzw. der Maskierung als Variable für einen einzelnen Buchstaben erfolgen⁵⁹. In der PLD wird die Suche nach Konjunktionen, Präpositionen, Interjektionen und ausdrücklichen Bestandteilen des Abfragevokabulars zwar nicht empfohlen, aber doch über die Verwendung von Anführungszeichen bzw. durch eine Zusammenfassung verwandter Wörter bei der Suche (z. B. „et“, „sed“, „ac“) durch das System nicht grundsätzlich von der Suche ausgeschlossen. Außerdem können unterschiedliche Schreibweise des Lateinischen durch eckige Klammern berücksichtigt werden⁶⁰. In THOMAS wird mit einem *Lemmatisierungsverfahren*, das die Wörter auf ihre lautliche bzw. semantische Wurzel verkürzt, die Integration aller grammatikalischen Varianten des eingegebenen Suchbegriffs in die Ergebnismenge ermöglicht, was eine Trunkierung nicht in dieser Vollständigkeit zu leisten vermag⁶¹. Es geht um einen semantisch-grammatikalischen, nicht nur

59) Bei ZID (hier nur Rechtstrunkierung; nützlich zur Unterscheidung von Einzelschlagwörtern und in Schlagwortketten verwendeten Schlagwörtern bei Expertensuche), RDB / SATB (Rechts- und Linkstrunkierung, Maskierung; außerdem Unterscheidung von exact und extended search, die eine automatische Rechtstrunkierung meint), MISSNORD (nur Rechtstrunkierung), Religio (Suche nach Wortbestandteilen statt exakter Suche entspricht einer Rechtstrunkierung), PLD (Rechts- und Linkstrunkierung, Maskierung), CETEDOC (Rechtstrunkierung und Maskierung), ACLL (Rechtstrunkierung), FRANCIS (Rechtstrunkierung).

60) Z. B. „epist[ou]la“

61) Bei der Lemmatisierung werden anders als bei der Trunkierung weder Homographen noch unterschiedliche Wortarten derselben Wurzel zusammengeführt (z. B. „magister“ als „Lehrer“ und als Ehrenbezeichnung für Petrus Lombardus; „possibilis“ und „possibilitas“ werden getrennt); dementsprechend ist eine Trunkierung nicht möglich.

formalen Abgleich zwischen Suchwort und Text. Aufwendige Kombinationsverfahren können vermieden werden. In der Ergebnisanzeige wird zwischen Lemma- und Form-Vorkommen unterschieden. In CETEDOC kann man die Funktion „formarum examen“ auswählen, wodurch alle Deklinations- und Konjugationsformen einbezogen werden⁶²). In ACLL wird eine noch weiter abgeschwächte Form der Induktion betrieben: die Enclitica (z. B. „-que, -ue, -ne, -met“) werden bei der Suche von dem Grundwort getrennt und mit einem davor gesetzten Bindestrich versehen, um eine höhere Treffermenge des Grundwortes in der allerdings jeweils konkreten grammatikalischen Form zu ermöglichen.

Bei THOMAS wird partiell eine *Deduktion in der Induktion* durchgeführt, wenn im Rahmen der Suche unter Nutzung des Lemmatisierungsverfahrens ein Auswahlmönü der verschiedenen grammatikalischen Wurzeln homographischer Suchwörter angeboten wird. Im ZID ist diese Funktion über den prioritären Registeranstieg mit seiner Übersicht über unterschiedliche Verwendungsweisen gleichlautender Schlagwörter (Homonymenzusätze, Verweisungen, Stammsatzerläuterungen) mitgegeben.

3. Korrelation von inhaltlicher Spezifität und technischer Multioptionalität

a) Indizes

Abgesehen von der Frage, inwieweit in der Indexgestaltung der formalen Konsistenz insbesondere der Autoren- und Personenschlagwortnamen entsprochen und über eine ausgefeilte Verweisungstechnik die vorhandene Variationsbreite der Sucheinstiege aufgefangen wird, ist auf zwei Charakteristika zu achten. Erstens besteht häufig eine *Korrespondenz des inhaltlichen bzw. methodischen Profils der Datenbank und der technischen Realisierung desselben*. So begegnet in FRANCIS ein vom „basic index“ zu unterscheidender „permuted index“, der die Schlagwörter aus unterschiedlichen Sprachen in einem Register zusammenfasst. PLD bietet einen eigenen Index für griechische Ausdrücke an, während in THOMAS gar keine Indizes verwendet werden, weil aufgrund der Lemmatisierung die Übernahme einzelner Indexeinträge sinnlos wäre. CETEDOC unterscheidet zwischen „titulus“ und „clavis“, wobei im Index zu letzterem Suchfeld nur diejenigen Titel eines Autors zu-

62) Da die Wörter wohl separat als Einzelwörter und als Bestandteile eines Wortfeldes mit den Flektionsformen erfasst wurden, ist hier eine Trunkierung möglich und auch sinnvoll, um z. B. eine wortartenübergreifende Suche zu betreiben.

sammengeordnet werden, die auf einer CD stehen, um eine Recherche oder Lektüre ohne fortwährenden Wechsel der CD zu ermöglichen⁶³⁾.

Zweitens hängt der Nutzen eines Index nicht nur von seinem Inhalt, sondern auch von dem Maß der *Beweglichkeit* und Markierungs- bzw. Übernahmefähigkeit der Einträge ab. Die Bewegung ist u. U. nur in einer Richtung möglich⁶⁴⁾ oder nur in umständlicher Weise⁶⁵⁾. Man kann häufig nur in begrenztem Maße mehrere Indexeinträge auf einmal übernehmen, was zumeist an der mangelnden Länge des betreffenden Suchfeldes liegt⁶⁶⁾. Im *ZID* kann hingegen bereits in kombinierter Form (z. B. als Schlagwortkette) gesucht und eine Kombination im Indexstatus - zudem über verschiedene Indizes hinweg - durchgeführt werden; für Aufruf und Bewegung bestehen jeweils mehrere Möglichkeiten⁶⁷⁾.

b) *Anzeigemodus*

Inhaltliche Eigenarten einer Datenbank sollten sich auch im Anzeigemodus widerspiegeln. Das kann etwa durch die exponierte oder wiederholte Auffüh-

63) Auch die bereits erwähnte Unterscheidung von „word“ und „entry“ beruht auf diesem Korrelationsverhältnis insofern, als dadurch die Nachteile einer starken Präkombination in der Sacherschließungsmethodik ausgeglichen werden sollen.

64) FRANCIS: Bewegung über „go-to“-Funktion, kein Zurückgelangen zu einer zuvor aufgerufenen Position ohne erneuten Indexaufruf; SATB: ein einmal eingegebener Eintrag innerhalb derselben Indexrecherche nicht revidierbar (anders z. B. bei ACLL, PLD).

65) MISSNORD: nur über „more“-Button; CETEDOC: nur über Scrollbalken, nicht über Cursor.

66) FRANCIS: keine längeren Kombinationen möglich, „or“-Verknüpfung bei Übernahme aus Index vorgegeben; RDB: zwar mehrere Indexeinträge übernehmbar, aber begrenzte Suchfeldlänge; SATB: Übernahme mehrerer Einträge (mit vorgegebener „or“-Verknüpfung) nur nach jeweils erneutem Indexaufruf; MISSNORD: Markierung und Übernahme nur dann möglich, wenn Indexeintrag bloß mit einem einzelnen Treffer verbunden ist; dagegen können in PLD mehrere Einträge in einem Zug markiert bzw. entmarkiert (in Ergebnislisten) werden und in CETEDOC sind bis zu hundert Einträge übernehmbar.

67) In der DOS-Version geschieht die Recherche stets über die Indizes, was Fehltreffer infolge defizitärer Eingabe verhindert. In Win-ZID kann ein zweites Indexfenster aufgerufen und zudem die Maus verwendet werden; allerdings muss hier die Kombination über die Expertensuche erfolgen. Als Proprium kann die Alternative von Kettendarstellung und Anzeige als Einzelelemente (Einzeldeskriptoren, Autor ohne Ausdifferenzierung der Werke, Zeitschrift ohne Einzelhefte) angesehen werden; dadurch wird die Ergebnismengenbildung (auch durch Übernahme in ein eigenes Ergebnismengenfenster, das Kombinationsmöglichkeiten mit logischen Operatoren bietet) erleichtert.

rung einzelner Kategorien geschehen⁶⁸⁾. Im Hinblick auf die weitere Verwendung des Datenmaterials ist bei Referenzdatenbanken die Unterscheidung einer reduzierten oder sämtliche Kategorien aufführenden⁶⁹⁾ bzw. einer kategorisierten, die Kategorien verbalisierenden sowie einer fortlaufenden Anzeigeform im ISBD-Format⁷⁰⁾ wichtig. Dem entspricht in Volltextdatenbanken die Wahlmöglichkeit zwischen satz- bzw. zeilenweiser und kontinuierlicher Ausführung des Textes, womit der deduktiv bzw. induktiv hierarchisierende Sacherschließungsmodus in der Anzeigeform manifest wird⁷¹⁾.

Die Arbeit mit den Anzeigebildschirmen wird durch *technische Optionen* wie Markierung und Texteinfügung⁷²⁾, Veränderung von Schriftgröße, -art und -farbe⁷³⁾ und Voreinstellung von Tastaturbelegungen oder Sortiermodus für Ergebnismengen erleichtert⁷⁴⁾. Ein Problem ist die Tatsache, dass je nach aktu-

-
- 68) In FRANCIS wird das Erscheinungsjahr, Unterscheidungskriterium für die Benutzung der einzelnen Teildatenbanken, zuerst an hervorgehobener Stelle und später noch einmal aufgeführt (Mitberücksichtigung der Kategorie „Sprache“ ebenfalls typisch ähnlich wie bei RDB); in SATB werden Erscheinungsjahr und Sprache exponiert positioniert, da der Unterscheidung von Englisch oder Afrikaans sowie der Zeit angesichts einer großen zeitlichen Ausdehnung für eine schnelle Relevanzentscheidung ein großes Gewicht zukommt; im ZID wird wegen der Ausrichtung auf die Bestände einer Bibliothek und um der Verknüpfung zum Dokumentlieferdienst (in Win-ZID eigener Flip zu einem Bestellformular) willen die Signatur der UB Tübingen auf einer eigenen Zeile aufgeführt.
- 69) Reduktionsmöglichkeit für die Anzeige: FRANCIS, RDB, SATB, MISSNORD (hier ist aber zwischen „standard“ und „long display“ kein Unterschied erkennbar; bei „brief display“ nur Hinweis auf eine Nummer innerhalb einer Gesamtmenge).
- 70) ZID: DOS: nur Kategorien- und bibliographisches Format; Windows: zusätzlich Format mit Verbalisierung der Kategorien; Reduktion nur durch Markierung für Export möglich.
- 71) PLD: über Symboltasten beliebige Hierarchisierung möglich; THOMAS: „inner hypertext“ und „plain text“ (mit allerdings recht kurzen Abschnitten); CETEDOC: „sententiae“ und „textus“ (dies fehlt bei ACLL).
- 72) Die Copy-and-Paste-Funktion ist an sich ein Proprium von Windows-Datenbanken, wird aber auch in der SATB angeboten. Dort kann man zudem ein „clipboard“, sozusagen eine beschreibbare Tafel als Erinnerungsstütze für die Entwicklung einer Recherchestrategie, benutzen; Teile aus Datensatzanzeigen können dorthin kopiert, Dateien geladen werden.
- 73) Die Schriftgröße und -farbe kann, obwohl eher Windows-spezifisch, in der DOS-Version des ZID über entsprechende Einträge in der Konfigurationsdatei verändert werden (anders z. B. ACLL).
- 74) So kann z. B. in RDB und SATB als Voreinstellung zwischen Standard- und Thesaurussuche, exakter und erweiterter Suche, kürzerer oder erweiterter Form der Kurzlistendarstellung, jedem Suchfeld als Sortierkriterium für die Ergebnislistendarstellung, auf- oder absteigender Auflistung, unterschiedlichen Farbkombinationen von Such- und Anzeigemaske gewählt und zudem die Belegung der Leer- und Entertaste bestimmt werden.

ellem Bildschirm die Betätigung derselben Funktionstaste andere Funktionen auslösen kann bzw. nicht alle Icons jeweils aktiviert sind.

c) *Datenexport*

Eine technische Multioptionalität im Hinblick auf den Datenexport liegt bei Referenzdatenbanken vor, wenn verschiedene Formate in Entsprechung zum Anzeigemodus sowie ein korrekter Transfer von Sonderzeichen angeboten wird. Dabei kann es zur Ausbildung eines *negativen Spezifikums* kommen, wenn der durch die Datenbanksoftware vorgegebene technische Rahmen nicht an die inhaltlichen Gegebenheiten und Erfordernisse der konkreten Datenbank angepasst wird und es zu redundanten Optionen kommt⁷⁵). Diesem Anliegen der *Vollständigkeit und gleichzeitigen Variabilität* entspricht bei Volltextdatenbanken die Frage, inwieweit der Kontext der Treffer und mögliche Querverbindungen wie Fußnoten mit exportiert werden können. Damit hängt das Problem der quantitativen Beschränkungen des Exports zusammen, die sowohl in einem Mangel an Speicherkapazität als auch in urheber- und wettbewerbsrechtlichen Motiven begründet sein dürften⁷⁶).

75) So wird in RDB und SATB eine für das vorliegende Datenmaterial unangemessene Kolumnendarstellung (Lotus, dBasell) oder ein Format „fixed fields“ zur Auswahl gestellt, bei dem die Kategorien und deren konkrete Füllungen nacheinander ausgegeben werden. Auch ist eine Bezugnahme auf Bilder im Exportfunktionsmenü überflüssig, weil in einer Referenzdatenbank keine nonverbalen Darstellungsformen vorgesehen sind. In FRANCIS kann nur mit ASCII exportiert werden. Im ZID ist hingegen ein Export im kategorisierten und im bibliographischen Format, in ASCII, RTF sowie in einem von WordPerfect interpretierbaren Format ebenso wie ein Export mitsamt allen Sonderzeichen möglich.

76) In PLD ist zwar nicht ein ganzer Band, wohl aber ein ganzes Dokument ohne den Gesamtapparat in einem Zug markier- und ausdrückbar. Das bei den Volltextdatenbanken durchgängige ASCII-Format reicht für das ohne Sonderzeichen auskommende Latein aus; griechische Wörter können in PLD mit PLD Greek und einem parallel laufenden Textverarbeitungsprogramm korrekt abgespeichert und ausgedruckt werden. In THOMAS und MISSNORD kann nur eine Texteinheit bzw. ein Datensatz gleichzeitig exportiert werden, in ACLL dagegen zwar eine große Menge an Einzeltreffern (formae) samt unmittelbarem Kontext, jedoch als fortlaufender Text nur recht kleine Texte. In CETEDOC können mehr als 30 Sätze nur dann ausgegeben werden, wenn sie nicht aufeinanderfolgen. In der RDB liegt die Begrenzung bei hundert Datensätzen; in FRANCIS und ZID ist eine Begrenzung allenfalls durch das verwendete Speichermedium (z. B. Diskette) vorgegeben. In den meisten Datenbanken kann die Suchstrategie mit abgespeichert sowie der Dateiname und -pfad der Exportdatei geändert werden.

d) *Hilfetexte*

Die einer Datenbank beigegebenen Hilfetexte sollten *übersichtlich* aufgeführt, untereinander *verbunden* und möglichst *mehrsprachig* angeboten werden⁷⁷⁾. Zudem sollten sie an die inhaltlichen Spezifika der konkreten Datenbank in ihren Bezugspunkten und Beispielen angepasst sein und sowohl primäre als auch komplexe, nicht in den Bereich der Selbsterklärung der Datenbank fallende Sachverhalte erläutern. Der Benutzer sollte ohne datenbankexterne Informationsquellen wie z. B. Handbücher auskommen können⁷⁸⁾.

4. Perspektiven oder: das Problem der Kompatibilität und Kooperation

In der Literatur wird als ein entscheidendes Beurteilungskriterium für Datenbanken deren *Konsistenz* im Hinblick etwa auf Namensansetzungen und Erschließungstiefe genannt⁷⁹⁾. Ein bleibendes Problem bei der Erlangung der Konsistenz ist allerdings ein gewisser Subjektivismus in der Sacherschließung, da bereits im kognitiven Akt der Inhaltsanalyse Vorverständnisse einwirken und sich diese in der Auswahl bzw. manchmal redundanten Neuanset-

77) Reine Online-Hilfen ohne thematische Übersichten liegen bei FRANCIS, ACLL und SATB vor. Eine Liste mit thematischen Unterblöcken und sowohl assoziativen als auch subordinativen Links bietet PLD. THOMAS arbeitet mit einem tastenbasierten Hyperlinksystem, während die Querverweise zwischen den Einzelthemen in CETEDOC per Mausclick aufrufbar sind. In der RDB werden in den Querverweisen und der Gesamtübersicht irritierenderweise unterschiedliche Bezeichnungen für dieselben Texte verwendet. ZID bietet nicht nur wie ACLL die Hilfen in mehr als einer Sprache an, sondern hat in Ergänzung zu den auf bestimmte Bildschirme und Funktionen bezogenen Online-Hilfen ausführliche und untereinander verknüpfte Hilfethemen in eigene Register integriert, wobei die Simultaneität der Versionen sich in Homonymenzusätzen zur Bezeichnung der DOS- bzw. Windows-Version widerspiegelt.

78) FRANCIS führt nirgendwo eine Verbalisierung der Klassifikation an. ACLL informiert nicht über die tatsächlich vorhandenen Begrenzungen der Exportmöglichkeit; das Handbuch ist notwendig, um die Zuordnung von Autoren und Werken zu eruieren. RDB übernimmt aus der vorgegebenen Software Beispiele für Patentinformations-, Telefon- und Adresdatenbanken.

79) So bei Ralph Köhler: Zeitschriftenaufsatz-Datenbanken Theologie: ein kritischer Vergleich zwischen Zeitschrifteninhaltsdienst Theologie und Religion Database unter Berücksichtigung weiterer Datenbanken. - Halle, 1998 (Schriften zum Bibliotheks- und Büchereiwesen in Sachsen-Anhalt ; 73), S. 15.18.30.74f.79; Petra Werner: „Dokumentation und Geisteswissenschaften“: zu Geschichte und aktuellen Problemen der Zeitschrifteninhaltserschließung - dargestellt anhand des Philosophischen Informationssystems (PHILIS) und des Zeitschrifteninhaltsdienstes Theologie (ZID). - Köln: Assessorarbeit, FH Köln, FB Bibliotheks- und Informationswesen, 1998, S. 125.

zung der sprachlichen Explikation der Inhalte in Form von Deskriptoren auswirken können⁸⁰). Außerdem ist insbesondere in Referenzdatenbanken eine Schichtung infolge der historischen Genese festzustellen, die insofern verkräftbar ist, als der Akzent auf der aktuellen Information über neuere Literatur liegt⁸¹.

Andererseits besteht das Anliegen einer möglichst weitgehenden *Vollständigkeit* des Materials und Konstanz der Erschließung, weil in den Geisteswissenschaften früher für unwichtig oder mittlerweile für veraltet Gehaltenes später doch an Bedeutung gewinnen kann⁸²). Insbesondere für die retrospektive Erschließung wird eine *Kooperation* der bisher eher autark wirkenden Datenbankproduzenten gefordert. Ein Kapazitätsgewinn sei durch den Abbau von Überlappungen und den Ausbau von Synergieeffekten zu erzielen⁸³). Wenn eine multilinguale Sacherschließung, die Verknüpfung der Schlagwörter mit den SWD-Identnummern, die Übernahme der Zeitschriftentitelaufnahmen aus der ZDB und die gerade im Hinblick auf eine Koppelung mit der Dokumentlieferung notwendige Eindeutigkeit der Angaben (ISBN, Paginierung) angestrebt wird⁸⁴), so geht es um die wechselseitige *Kompatibilität* der Datenbankstrukturen. Davon abgesehen, dass eine Akzeptanz der deutschen Normierungen im Ausland nicht anzunehmen ist, würde ein Zusammenspielen der fachbezogenen Daten unterschiedlichster Datenbanken, der kirchlichen Bibliotheken

80) Vgl. dazu Holger Nohr: „Inhaltsanalyse“. - In: NfD 50 (1999), S. 71. Das Problem eines Subjektivismus in der sprachlichen Wiedergabe von Sachverhalten wirkt sich aber ebenfalls in Volltextdatenbanken aus, weil sowohl die Autoren der Texte als auch die Rechercheure davon betroffen sind und eine u. U. divergierende Terminologie mit bestimmten Inhalten assoziieren. Das Normierungssystem mit der Festlegung von Vorzugsbenennungen und Verweisungen versucht diese Schwierigkeiten abzumildern.

81) Zum evolutiven Charakter und der tendenziellen Current-Contents-Funktion Werner, Dokumentation, S. 111f.117. Bernhard Schmitt: „Die CD-ROM-Datenbanken des Bischöflichen Priesterseminars Trier“. - In: Mitteilungsblatt der Arbeitsgemeinschaft Katholisch-Theologischer Bibliotheken 43 (1996), S. 151-163; ebd., S. 154 bzgl. CETEDOC (Verifikationsarbeiten, Flüchtigkeitsfehler): „opus semper perfectum“.

82) Vgl. Werner, Dokumentation, S. 110 (hier auch Forderung der Transparenz bzgl. Auswahlkriterien und ausgewähltem Material).129.143; Köhler, Zeitschriftenaufsatz-Datenbanken, S. 38.

83) So bei Köhler, Zeitschriftenaufsatz-Datenbanken, S. 5.35f.43; Werner, Dokumentation, S. 121.

84) Vgl. Köhler, Zeitschriftenaufsatz-Datenbanken, S. 15.18.22.31.44.50; zu der im ZID durch seine Ausrichtung auf die physischen Bestände einer Bibliothek schon von Anfang an vorhandenen Verbindung mit dem Dokumentlieferservice SSG-S: Werner, Dokumentation, S. 137.139.

und Verbände in einen gemeinsamen Datenpool oder auch die Fusion von ZID und RDB gerade die innere Konsistenz zerstören, die zuvor als Voraussetzung für Kooperationsmaßnahmen intendiert worden war. Die dann nur konsequente Forderung nach einem Verzicht auf jegliche Sacherschließung zugunsten einer reinen Stichwortsuche in einem zwar vollständigeren, aber auch unübersichtlicheren Datenreservoir verkennt die Funktion einer Datenbank, eine Filterung der Informationen auf Relevanz hin zu unterstützen⁸⁵⁾; insofern gilt: „weniger ist mehr“⁸⁶⁾.



85) Vgl. Köhler, Zeitschriftenaufsatz-Datenbanken, S. 46.47f.50; Werner, Dokumentation, S. 130, bemerkt richtig, dass zwar ein Scannen der Volltexte und eine Zusammenführung der Volltexte mit den Referenzdatenbanken wünschenswert sein könnte, dieses Verfahren aber gerade nicht die formale und sachliche Erschließung überflüssig macht, weil die Unterscheidung des Wichtigen und Unwichtigen in den Volltexten unterstützt werden muss. Das Aktualitätsproblem ließe sich zwar vordergründig durch ein Angebot im Internet lösen; dieser Gewinn würde allerdings durch eine erhebliche Verlangsamung des Retrievals erkauft werden (Werner, ebd., S. 112, Anm. 29).

86) Werner, Dokumentation, S. 122. Grundsätzlich richtig ist sicherlich die Idee einer stärkeren Offenheit der Referenzdatenbanken für Informationen über bereits erfolgte Digitalisierungen der Volltexte oder eine Integration von Nachschlagewerken (Werner, ebd., S. 129.130.137). Mit Daten, die nicht nur in ihrem Wert, sondern auch in ihrer Richtigkeit nach kurzer Zeit überholt sind (z. B. Hinweise auf laufende wissenschaftliche Arbeiten zu bestimmten Themen oder Veranstaltungen; so bei Werner, ebd., S. 131), sollte man eine Datenbank allerdings um des Arbeitsaufwandes und Niveauunterschiedes der Information willen besser nicht überfrachten. Bei der Kooperationsthematik wird außer der Differenz in der methodischen Konzeption zudem die gegensätzliche ökonomische Intention ausgeklammert. Die Preisunterschiede zwischen Datenbanken, die mit kommerziellen Interessen und auch Notwendigkeiten produziert werden, und solchen, die sich als Dienstleistung von Non-Profit-Organisationen verstehen, sind beträchtlich (RDB: \$ 1835 plus \$ 300 für Netzzugang / ZID: DM 350; SATB nur DM 655 wegen ausgeprägten Sponsorings; PLD: £ 21600 / THOMAS: DM 1700 (allerdings auch wegen unterschiedlichen Textvolumens und anderer Software).