Aufsätze

Der Kooperative Bibliotheksverbund Berlin-Brandenburg

von Martin Grötschel • Monika Kuberek • Stefan Lohrum • Joachim Lügger • Beate Rusch

Im Herbst 1999 ging der Kooperative Bibliotheksverbund Berlin-Brandenburg (KOBV) in Routinebetrieb. In technischer Hinsicht setzt der KOBV auf ein neuartiges, zukunftsweisendes Verbundkonzept, in dem die zentrale Datenbank mit dem Verbundkatalog ersetzt wird durch verteilte Datenbanken und lokale Kataloge, die durch das Netz (Internet) miteinander verbunden sind. Die zentrale Rechenfunktion wird ersetzt durch die KOBV-Suchmaschine. Beschrieben werden die Hintergründe, die zu diesem Konzept geführt haben, die Kriterien, die die Architekturentscheidungen der Software bestimmt haben, sowie die einzelnen realisierten Komponeneten.

The Kooperative Bibliotheksverbund Berlin-Brandenburg (KOBV) took up routine operation in Autumn 1999. From the technical point of view KOBV initiates a new forward-looking union concept which foresees the central database being replace by the union catalogue through its distributed data base system and local catalogues, which are networked (Internet). The retrieval functions will be carried out by the KOBV search engine instead of a central database functionality. The backround to this concept, the criteria for dererming the software architecture and the individually realised components are described here.

Inhaltsübersicht

1 Einleitung
2 Die Ausgangssituation
3 Die Empfehlungen für die Region
4 Das KOBV-Projekt am Konrad-Zuse-Zentrum für Informationstechnik Berlin
5 Lokale Systeme
6 Realisierung der Suchmaschine: Ausschreibung
7 Architekturentscheidungen
7.1 Offenheit des Systems
7.2 Performanz des Systems
7.3 Skalierbarkeit des Systems
7.4 Komfortable Recherchemöglichkeiten
7.5 Systemarchitektur: Quick Search in der Verteilten Suche
8 Parallele Verteilte Suche: Zugriff mit WWW-Browsern
8.1 Die WWW-Oberfläche der KOBV-Suchmaschine
8.2 Verzweigen in die lokalen OPACS (WEB-Link)
9 Sequentielle Verteilte Suche: Zugriff mit Z39.50 Clients
10 Gemeinsamer Index (Quick Search)
10.1 Indexinformationen - Datenstruktur
10.2 Dublettenbehandlung
10.3 Update
11 Ausblick
Anhang
Abkürzungsverzeichnis
Abbildungsverzeichnis

1. Einleitung

Im Herbst 1999 stellte sich der Kooperative Bibliotheksverbund Berlin-Brandenburg (KOBV) der Öffentlichkeit mit seinem neuen Verbundkonzept vor. Technisches Herzstück dieses Verbundes ist die KOBV-Suchmaschine, die an die Stelle des herkömmlichen zentralen Verbundkataloges tritt. Die KOBV-Suchmaschine sucht die verschiedenen Kataloge der Region unter einer einheitlichen Oberfläche ab und leitet die Benutzer über einen Link in die lokalen Opacs weiter. Dort kann der Benutzer die gefundenen Dokumente bestellen oder vormerken lassen. Unter der WWW-Adresse *http://www.kobv.de* wird neben dieser verteilten Suche auch eine Schnelle Suche (Quick-Search) angeboten, die ein zentraler Index ermöglicht. Gleichzeitig ist die KOBV-Suchmaschine als Hilfe für die Katalogisierung gedacht. Der Bibliothekar recherchiert über die KOBV-Suchmaschine aus seinem lokalen Bibliothekssystem heraus in den verschiedenen Katalogen und übernimmt Daten in die eigene lokale Datenbank.
Dieser Artikel stellt die Hintergründe dar, die zu diesem bislang einmaligen Verbundsystem führten. Beschrieben werden die Gesichtspunkte, die bei der Realisierung der Suchmaschine zu berücksichtigen waren, ebenso wie die entwickelten Funktionalitäten im einzelnen. Dieser Aufsatz versteht sich als eine Art Status-Bericht mit Stand Oktober 1999. Einige der in diesem Aufsatz erst für die Zukunft angekündigten Weiterentwicklungen werden vermutlich bereits bei Erscheinen dieses Heftes umgesetzt worden sein.

2. Die Ausgangssituation

Der Bibliotheksverbund Berlin-Brandenburg (BVBB), der seine Anfänge in den 80er Jahren hat, betrieb keinen eigenen Rechner, sondern nutzte für seine Verbundanwendung den Großrechner des Deutschen Bibliotheksinstituts (DBI). Das DBI als Anbieter der Zeitschriftendatenbank sowie die Länder Bayern, Baden-Württemberg und Nordrhein-Westfalen entschieden sich Mitte der 90er Jahre dafür, ihre Großrechner-Verbundsysteme gemeinsam abzulösen und das System Horizon der Firma Dynix einzusetzen. Der BVBB schloß sich diesem Verbundkonsortium nicht an.

Anschrift der Verfasser:
Martin Grötschel • Monika Kuberek • Stefan Lohrum • Joachim Lügger • Beate Rusch, Konrad Tuse Institut für Informationstechnik, Takustr. 7, 14195 Berlin

Die Zeit drängte, zumal das im DBI eingesetzte Betriebssystem BS2000 nicht Jahr-2000-fähig war. In dieser Situation beauftragten die zuständige Berliner Senatsverwaltung für Wissenschaft, Forschung und Kultur und das gleichnamige Ministerium des Landes Brandenburg eine auswärtige Gruppe von Experten, Empfehlungen für die Region zu erarbeiten.[1] Die Experten, deren Empfehlungen für den Kooperativen Bibliotheksverbund Berlin-Brandenburg letztendlich angenommen wurden, waren B. Dugall (Leitender Bibliotheksdirektor der Stadt- und Universitätsbibliothek Frankfurt am Main), M. Hebgen (Abteilungsleiter im Universitätsrechenzentrums Heidelberg) sowie Prof. Dr. W. König (Institut für Wirtschaftsinformatik am Fachbereich Wirtschaftswissenschaften der Universität Frankfurt).

3. Die Empfehlungen für die Region

Die Experten gaben der an Beständen reichen Bibliotheksregion Berlin-Brandenburg keine gute Noten. Die Leistungen des bestehenden Verbundes BVBB stünden in keinem vertretbaren Kosten-Nutzen-Verhältnis, so lautete pointiert ihre Kritik. Die Nachweissituation in der Region war denkbar schlecht: Nur ein kleiner Teil der wissenschaftlichen Bestände war im BVBB nachgewiesen, der größere Teil im Zettelkatalog des Berliner Gesamtkataloges (BGK), dessen Pflege mit dem Jahr 1991 eingestellt wurde, große Bestände zudem nur vor Ort, in den Zettelkatalogen der Bibliotheken.
Die Experten stellten die personalintensive Pflege eines zentralen Verbundkataloges wie des BVBB grundsätzlich in Frage. Vor allem weil dieser vom Benutzer nur als Mikrofiche-Katalog bzw. als neu ausgedruckter Eintrag in einem der zahlreichen Zettelkataloge wahrgenommen werde. In Frage stellten sie in letzter Konsequenz auch die Regeln und Konventionen, nach denen die Bibliothekare nicht nur in dieser Region seit den 70er Jahren arbeiteten. Diese Kritik betraf die hochspezialisierten Regelwerke ebenso wie das bibliothekarische Austauschformat MAB (Maschinelles Austauschformat für Bibliotheken).[2]

Die Experten beschrieben modellhaft sechs alternative Systemarchitekturen für Bibliotheksverbünde:
- Modell 1: Vollkommen zentralistisches System am Beispiel von ETHICSplus (ETH - Zürich)
- Modell 2: Zentralistisches System mit top-down-getriebenen lokalen Komponenten am Beispiel von PICA (Deutschland)
- Modell 3: Bottom-up-getriebenes zentrales System mit lokalen Komponenten am Beispiel von OhioLink
- Modell 4: Konföderiertes Servernetz mit Union Catalog
- Modell 5: Konföderiertes Servernetz ohne Union Catalog, aber mit einem einheitlichen Fremddatenserver
- Modell 6: Konföderiertes Servernetz ohne Union Catalog und mit verteilten Fremddatenservern

Die exemplarisch beschriebenen Systemarchitekturen unterscheiden sich durch ihre unterschiedlich starke Gewichtung der lokalen Komponenten (= lokale Systeme). Die Spannbreite der Möglichkeiten reicht von dem mit EthicsPlus beschriebenen System, in dem alle Informationen zentral gespeichert werden und das als Eingabegeräte in den teilnehmenden Bibliotheken nur „dumme Terminals" benötigt (Modell 1) bis zu einer dezentralen Architektur ohne jede zentrale Komponente, die mit jedem lokalen System einen weiteren autonomen Knoten im Netz bekommt (Modell 4). Die Modelle 5 und 6 gehen gesondert der Frage nach der zentralen oder dezentralen Fremddatenhaltung nach. Dabei kennt Modell 6 keinerlei zentrale Datenhaltung mehr und verteilt auch das Fremddatenangebot auf lokale Server.

Für die Region schlugen die Experten das Modell 5 vor. Als zentrale Datenbank des Verbundes bleibt danach nur der Fremddatenserver bestehen. Abgeschafft werden soll dagegen der zentrale Verbundkatalog zugunsten einer starken Betonung der lokalen Komponenten. Abgeschafft werden sollen damit auch die zum Betrieb und zur zentralen Datenpflege benötigten Personalkapazitäten in der Verbundzentrale. Die Funktionen, die bislang der Verbundkatalog übernimmt, den Nachweis über die im Verbund vorhandenen Bestände, soll eine Suchmaschine leisten.

Die Experten empfahlen der Region:
- Die Abschaffung des zentralen traditionellen Kataloges
- Die Suche über heterogene Systeme unterschiedlicher Hersteller über internationale, standardisierte Schnittstellen (Z39.50/WWW)
- Einen Minimalstandard für die Katalogisierung angelehnt an die Elemente des Dublin Core

Damit wurde für die berlin-brandenburgische Bibliothekslandschaft eine Systemarchitektur vorgeschlagen, die sich in verschiedenen Wirtschaftsbereichen bereits bewährt hatte. Im Bibliothekswesen allerdings ist die Idee einer hochgradig verteilten Struktur gleichberechtigt nebeneinanderstehender Server, d.h. lokaler Systeme, die ohne eine zentrale Primärdatenbank nur noch über das Netz miteinander verbunden sind, nach wie vor neu.

4. Das KOBV-Projekt am Konrad-Zuse-Zentrum für Informationstechnik Berlin

Die Empfehlungen von Dugall/Hebgen/König stießen in der deutschen Bibliothekslandschaft auf große Skepsis. Was vor allem vermißt wurde, war die Einbettung der Bibliotheksregion Berlin-Brandenburg in das nationale Gefüge bibliothekarischer Kooperationen und Konventionen. Lehnte man in Berlin und Brandenburg den regionalen Alleingang in Richtung US-MARC und AACR - so wie in den Empfehlungen vorgeschlagen - ab, so sah man außerhalb der Region die Gefahr, der neue Verbund werde sich bei den übrigen, zentral strukturierten Verbünden und seinen qualitativ hochwertigen, von Zentralredaktionen gepflegten Daten „bedienen", ohne selber entsprechende Daten zu liefern.[3]
Bibliothekare und Informatiker aus den Berliner und Brandenburger Bibliotheken diskutierten in der Folge die Empfehlungen und arbeiteten konkrete Umsetzungskonzepte aus.[4] Nach diesen ersten konzeptuellen Festlegungen entschlossen sich die Länder, die Empfehlungen in einem wissenschaftlichen Projekt mit zunächst dreijähriger Laufzeit umzusetzen. So wurde am 1. April 1997 der Kooperative Bibliotheksverbund Berlin-Brandenburg (KOBV) als Projekt begründet. Angesiedelt ist das Projekt am Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB). Auch mit der Entscheidung, mit der Realisierung des neuen Verbundes eine Einrichtung außerhalb des Bibliothekswesens zu betrauen, folgten die Länder Berlin und Brandenburg den Empfehlungen des auswärtigen Experten.
Das Konrad-Zuse-Zentrum für Informationstechnik Berlin ist eine außeruniversitäre Forschungseinrichtung des Landes Berlin. Es betreibt Forschung und Entwicklung auf dem Gebiet der Informationstechnik, vorzugsweise in anwendungsorientierter algorithmischer Mathematik. Zugleich bietet es Höchstleistungsrechnerkapazität als dazugehörige Dienstleistung an.[5] Im Bereich der Informationsvermittlung hat das ZIB insbesondere mit dem in der Abteilung Wissenschaftliche Informationssysteme angesiedelten Math-Net-Projekt einschlägige Erfahrungen gesammelt.[6]

In einer gemeinsamen Vereinbarung verständigten sich die Länder und das ZIB auf die folgenden 15 Projektpartner [7]:

- Bibliothekssystem der Freien Universität Berlin
- Universitätsbibliothek der Humboldt-Universität zu Berlin
- Universitätsbibliothek der Technischen Universität Berlin
- Bibliothek der Hochschule der Künste Berlin
- Universitätsbibliothek der Brandenburgischen Technischen Universität Cottbus
- Universitätsbibliothek der Europa-Universität Viadrina Frankfurt (Oder)
- Universitätsbibliothek der Universität Potsdam
- Hochschulbibliothek der Hochschule für Film und Fernsehen „Konrad Wolf", Potsdam
- Bibliothek der Fachhochschule Brandenburg
- Bibliothek der Fachhochschule Eberswalde
- Bibliothek der Fachhochschule Lausitz, Senftenberg
- Bibliothek der Fachhochschule Potsdam
- Bibliothek der Technischen Fachhochschule Wildau
- Stadt- und Landesbibliothek Potsdam
- Staatsbibliothek zu Berlin - Preußischer Kulturbesitz

Das auf drei Jahre angelegte Projekt Kooperativer Bibliotheksverbund wurde Mitte des Jahres 1999 um neun Monate bis Ende 2000 verlängert. In dieser Zeit sollen von den Trägern die Grundlagen für die Institutionalisierung des KOBV ab 2001 geschaffen werden. Dem Projekt stehen vier wissenschaftliche Mitarbeiter (zu gleichen Teilen Informatiker und Bibliothekare) sowie eine Sekretärin zur Verfügung. Diese Projektgruppe ist dem Abteilungsleiter Wissenschaftliche Informationssysteme, J. Lügger, zugeordnet. Die Gesamtprojektleitung hat der Vizepräsident des ZIB, Professor Dr. M. Grötschel, inne.

5. Lokale Systeme

Die Empfehlungen der Gutachter flossen im Frühjahr 1997 in einen Entwurf für ein Gemeinsames Landeskonzept der Länder Berlin und Brandenburg ein, welches für das Land Berlin mit dem Beschluß des Abgeordnetenhauses vom 8. Dezember 1998 auf landespolitischer Ebene bestätigt wurde.[8]
Die Suchmaschine, die in dem Gemeinsamen Landeskonzept als zentrale Verbundkomponente für die Region verbindlichen Charakter bekam, setzt die Existenz von modernen lokalen, integrierten Bibliothekssystemen bei den Verbundteilnehmern voraus. Für welches lokale System sich eine Bibliothek entscheidet, spielt dabei keine Rolle, solange bestimmte Schnittstellen vorhanden sind.

Es wird vorausgesetzt, daß der jeweilige Bestand lokal sowohl über das WWW (WWW-OPAC) als auch über Z39.50 (Z39.50-Server) zugänglich ist. Weiterhin wird zur Übernahme von Daten aus dem gemeinsamen Fremd- und Normdatenangebot sowie von Katalogisaten anderer Verbundbibliotheken lokal eine (integrierte) Z39.50-Client-Schnittstelle benötigt. Diese Schnittstellen sind bei Systemen älterer Provenienz nicht unbedingt implementiert.

Die politische Willensbekundung der Länder Berlin und Brandenburg für einen gemeinsamen Bibliotheksverbund schuf für die Berliner Hochschulen die (längst überfällige) Grundlage, auf der sie bei der Deutschen Forschungsgemeinschaft (DFG) durchsetzungsfähige Förderanträge zur Ausstattung ihrer bibliothekarischen Einrichtungen mit modernen Bibliotheksmanagementsystemen stellen konnten. In Brandenburg bedeutete die anstehende Realisierung des KOBV, von vorhandener EDV auf Systeme der nächsten Generation zu migrieren.

Die Folge war ein wahrer Modernisierungsschub in der Bibliotheksregion Berlin-Brandenburg.

- In Berlin wurde 1999 in den vier großen Hochschulbibliotheken der Technischen Universität, der Freien Universität, der Humboldt-Universität und der Hochschule der Künste das Bibliothekssystem *Aleph500* eingeführt.
- Die Staatsbibliothek entschloß sich, aktiv in den Gemeinsamen Bibliotheksverbund (GBV) mit Sitz in Göttingen zu katalogisieren und entsprechend lokal die *PICA*-Katalogisierungskomponente und den PICA-Opac einzusetzen. Allerdings unterzeichnete die Staatsbibliothek den Vertrag mit dem GBV unter der Vorgabe, ihren Bestand gleichzeitig in der Region über den KOBV recherchierbar machen zu können.
- Die Universitätsbibliotheken Frankfurt (Oder) und Cottbus sowie sämtliche brandenburgischen Fachhochschulbibliotheken migrierten im Laufe des Jahres 1999 von *SISIS-SE* auf die Version *SISIS 5.0*.
- An der Stadt- und Landesbibliothek Potsdam ist das UNIMARC-basierte *Horizon-System* des Anbieters *Dynix* im Einsatz.
- An der Universitätsbibliothek Potsdam besteht die Absicht, von *Allegro* auf ein neues System zu migrieren.
- In der Bibliothek der Hochschule für Film und Fernsehen Konrad-Wolf steht die Migration von *Bislok* auf Bibliotheca 2000 des Anbieters *BOND* bevor.

Damit sind heute unter den 15 Projektpartnern 6 verschiedene Systeme im Einsatz, die alle über WWW- und Z39.50-Schnittstellen verfügen. Bis auf das Allegro-System, für das die Client-Schnittstelle noch in Entwicklung begriffen ist, ist auch bei allen ei-

ne Z39.50-Client-Schnittstelle vorhanden. Alle genannten Bibliotheken stellen ihre Bestände über WWW zur Verfügung, wenngleich mit unterschiedlichem Funktionsumfang für die Benutzer.

6. Realisierung der Suchmaschine: Ausschreibung

Zu glauben, eine Suchmaschine mit dem Funktionsumfang eines Verbundinformationssystems, wie sie für den KOBV geplant ist, als Public Domain-Produkt im Internet zu finden, wäre naiv gewesen. Eine Eigenentwicklung wurde am ZIB kurzzeitig in Betracht gezogen, jedoch schnell wieder verworfen. Gegen eine Eigenentwicklung sprach hauptsächlich, daß das ZIB die zwingend geforderte Gewährleistung, Pflege des Produkts sowie dessen Weiterentwicklung nicht hätte garantieren können. Vor diesem Hintergrund wurde entschieden, die Suchmaschine auf der Basis eines am Markt befindlichen Bibliothekssystems zu realisieren. Man hoffte, einen Anbieter zu finden, für den die Realisierung der Suchmaschine eine strategische Investition in die Zukunft bedeutete und der sie in die eigene Versionsplanung aufnehmen würde.
1997 schrieben die vier Berliner Hochschulen, die Technische Universität Berlin (TU), die Freie Universität (FU), die Humboldt-Universität (HU) und die Hochschule der Künste (HdK), die Suchmaschine gemeinsam mit den lokalen Bibliotheksinformationssystemen als Paket aus. Das ZIB selbst beteiligte sich an dieser Ausschreibung nicht, sondern nahm nur beratende Funktion wahr. Nach einem aufwendigen Auswahlverfahren fiel die gemeinsame Entscheidung der auslobenden Einrichtungen auf das System Aleph500 des israelischen Anbieters Ex Libris Ltd. [9] Im Juli 1998 schlossen das ZIB und die genannten Hochschulen einen Konsortialvertrag mit der deutschen Tochter, Ex Libris Deutschland GmbH, der die Ausstattung der oben genannten Bibliotheken mit dem integrierten Bibliothekssystem Aleph500 sowie die Realisierung einer zentralen Suchmaschine zum Inhalt hat.

7. Architekturentscheidungen

Die in den Empfehlungen avisierte Suchmaschine sollte gleichermaßen die Interessen der Nutzer wie die der Bibliothekare berücksichtigen. Die Suchmaschine stellt damit ein komplexes Verbundinformationssystem dar, das dem Nutzer über das Internet einen schnellen Nachweis über die Bestände der Region (möglichst mit Bestellunterstützung) und dem Bibliothekar eine Hilfe zur Nutzung der Katalogisate aus den Partnerbibliotheken bieten will.[10]
Die Kriterien, die die Entscheidung für die nun realisierte Architektur der KOBV-Suchmaschine geleitet haben, werden im folgenden beschrieben.

7.1 Offenheit des Systems

Die grundsätzliche Entscheidung, mit dem KOBV einen Verbund zu realisieren, in dem das integrierte Arbeiten mit heterogenen Bibliothekssystemen auf lokaler Ebene möglich, wenn nicht gar gefördert werden soll, erhebt die Verwendung von internationalen Standards zu einer zwingenden Notwendigkeit. Die KOBV-Projektgruppe am ZIB sieht den Einsatz standardisierter Schnittstellen als Voraussetzung für die Kommunikation zwischen Systemen unterschiedlicher Provenienz. Dadurch lassen sich umfangreiche Anpassungsarbeiten seitens der lokalen Systeme an die Verbundsoftware vermeiden. Die Standardschnittstelle Z39.50 und die des Internets sind der Garant für die Offenheit der Verbundarchitektur.

Für den Bereich des Information Retrieval ist das offene Protokoll Z39.50 entwickelt worden.[11] Dieser ANSI/NISO-Standard definiert die Art und Weise, wie zwei Computer miteinander kommunizieren. Eversberg beschreibt das Z39.50-Protokoll treffend als Esperanto für Datenbanken.[12] Ebenso wie die Kunstsprache Esperanto die Verständigung von Sprechern verschiedener Muttersprachen möglich macht, abstrahiert auch Z39.50 von den verschiedenen herstellerspezifischen Retrievalsprachen und proprietären Schnittstellen. Unter der Voraussetzung, daß beide Seiten das „Datenbank-Esperanto Z39.50" verstehen, d.h. die Schnittstelle implementiert haben, wird so die Kommunikation zwischen Systemen verschiedener Hersteller möglich, unabhängig auch von deren Hardware und Betriebssystem.
Für den Benutzer liegt der Charme dieses Protokolls darin, daß er in die Lage versetzt wird, in einer verteilten Client-Server-Umgebung in Systemen unterschiedlicher Hersteller zu recherchieren, ohne deren spezifische Anfragesprachen erlernen zu müssen. Er kann die Retrievalsprache seines „Heimatsystems" verwenden. Das gilt für den Bibliotheksbenutzer ebenso wie für den Bibliothekar, der in fremden Systemen nach Datensätzen mit dem Ziel recherchiert, diese in das eigene System zu übernehmen.
Als 1998 die Entscheidungen für die Architektur der Suchmaschine fielen, befand sich die Implementierung des Z39.50-Protokolls bei den in Deutschland verbreiteten Bibliothekssystemen - im Unterschied etwa zu den USA - noch in den Anfängen. Zwar war mit dem DBV-OSI-Projekt die Schnittstelle auf Seiten der Verbundsysteme realisiert worden, die entsprechende Integration der Client-Schnittstelle in die Katalogisierungs- bzw. OPAC-Module der lokalen Systeme mußte jedoch erst noch geleistet werden.[13] Vor diesem Hintergrund entschied die KOBV-Projektgruppe bereits in der Phase der Spezifizierung, von den umfangreichen Möglichkeiten des Z39.50-Protokolls nur die Grundfunktionen zu nutzen und entsprechend auch bei den lokalen Systemen vorauszusetzen. Diese Funktionen - in der Sprache des Z39.50 Protokolls Services - beschränken sich auf Init (Verbindungsaufbau), Search (Suche), Present (Anzeige von Ergebnissen) und Scan (Registersuche).

Da eine komfortable Suche in den Katalogen der KOBV-Bibliotheken von jedem Arbeitsplatz aus möglich sein soll, wird als zweite Standardschnittstelle bei den lokalen Systemen wie auf Seiten der Suchmaschine eine WWW-Schnittstelle vorausgesetzt. Über die WWW-Schnittstelle der Suchmaschine kann mit jedem handelsüblichen Internet-Browser in den Beständen des KOBV recherchiert werden. Technisch setzt das die Realisierung eines WWW-Z39.50-Gateways voraus.[14]

Will man von der Suchmaschine in die lokalen Systeme über einen Hyperlink weiterschalten, muß auch hier eine WWW-Schnittstelle vorhanden sein.

7.2 Performanz des Systems

Die Performanz des Systems war und ist neben der verlangten Offenheit das ausschlaggebende Argument in der Diskussion über die Suchmaschinen-Architektur. Die Akzeptanz des Systems bei beiden Nutzergruppen - Bibliotheksbenutzern wie Bibliothekaren - steht und fällt mit dem Antwortzeitverhalten. Als Richtwert gilt hier, daß das Ergebnis einer Suche in 3-5 Sekunden angezeigt werden soll. Das Blättern in Registern und in Ergebnislisten stellt mit etwa 1-2 Sekunden allerdings erheblich höhere Anforderungen an die Performanz.

Für den Nutzer setzt sich die Antwortzeit aus folgenden Faktoren zusammen:
- Recherche in der Datenbank
- Übermittlung der Daten aus der Datenbank
- Darstellung der Daten im Client (WWW oder Z39.50-Client)

Das Antwortzeitverhalten der Suchmaschine ist neben der eigenen Leistungsfähigkeit von mehreren externen Faktoren abhängig:
- Geschwindigkeit / Auslastung der Netzverbindungen
- Antwortzeitverhalten / Leistung der Zielsysteme

Bei den Zielsystemen spielt dabei deren Netzanbindung sowie die Leistungsfähigkeit der lokalen Server eine Rolle. Die Leistung der lokalen Rechner wiederum ist abhängig von der jeweiligen Hardwareausstattung sowie deren Konfiguration. Erheblichen Einfluß auf die Antwortzeit hat auch die Größe der Ergebnismenge.[15]

Da die KOBV-Suchmaschine zukünftig auch kleinere Bibliotheken integrieren will, gilt allerdings grundsätzlich die Vorgabe, die Einstiegsschwelle möglichst niedrig zu halten und sie nicht durch die Forderung nach einer bestimmten Hardwareausstattung zu erhöhen.

7.3 Skalierbarkeit des Systems

Die Forderung nach einer hochgradigen Skalierbarkeit des Systems, ergibt sich nicht zuletzt aus der großen Anzahl der Bibliotheken in Berlin und Brandenburg. Um einen optimalen Nachweis in der Region zu schaffen, ist die Integration der vielen bibliothekarischen Einrichtungen, darunter auch kleine hoch spezialisierte Bibliotheken, in den KOBV fast zwingend. Langfristig kann das bedeuten, über die KOBV-Suchmaschine die Bestände von rund 1000 Bibliotheken verschiedenster Größe aus der Region nachzuweisen.

Die Anfragen der Suchmaschine darf allerdings die unterschiedlich leistungsfähigen lokalen Systeme nicht übermäßig belasten. In diesem Zusammenhang sind folgende Maßnahmen vorgesehen:
- Beschränkung der maximalen Anzahl der Bibliotheken, in denen ein Benutzer gleichzeitig suchen kann.
- Steuerung der Auswahl von Bibliotheken, in den recherchiert werden soll, durch benutzerdefinierte Suchprofile und virtuelle Fachbibliotheken.

Die Suchmaschine wurde für etwa 200-300 simultane WWW-Benutzer ausgelegt. Um diesen Anforderungen gerecht zu werden, wurde eine ausbaufähige Hardwareausstattung gewählt. Sie besteht derzeit aus einem Server SUN E3500 mit 4 Prozessoren und 3 GB Hauptspeicher sowie ein über Fibre-Channel verbundenes Disk-Array A5000 mit ca. 200 GB Plattenplatz. Ein technisch ähnlich ausgestattetes System dient als Testsystem und der Ausfallsicherung.

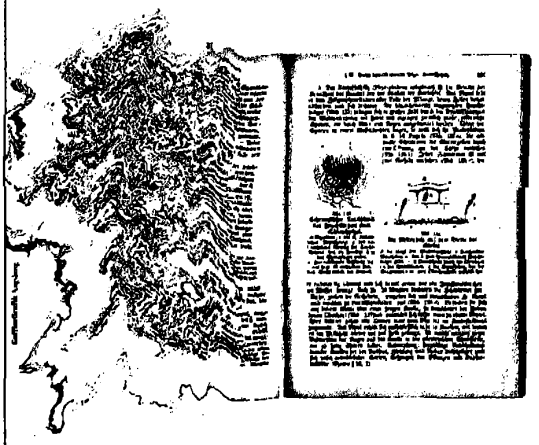
7.4 Komfortable Recherchemöglichkeiten

Zu einer komfortablen Suche gehört neben der Möglichkeit, mit Einzelwörtern bzw. Phrasen unter zuhilfenahme Boolescher Verknüpfungen zu suchen, die Suche über alphabetische Register. Dieser Sucheinstieg gewinnt besonders an Bedeutung in einer heterogenen Umgebung, in der man mit den verschiedenen Zielsystemen nicht vertraut ist und weder weiß, welche Titel verzeichnet sind, noch welche Wörter wie indexiert sind. Die technische Realisierung von alphabetischen Registern stellt für die verteilte Suche eine besondere Herausforderung dar. Einerseits darf von unterschiedlichen Indexierungspraktiken in den Bibliotheken ausgegangen werden (Indexierung als Wort oder als Phrasenregister, Einzelregister, Kreuzregister). Andererseits muß dieser Sucheinstieg in der geforderten Geschwindigkeit (performant) erfolgen.

Hinsichtlich der erzielten Ergebnisse sind in einer verteilten

Säure im Papier zerstört im Laufe der Zeit wertvolle Bücher.

Das einmalige, patentierte Datukom-System sorgt für eine vollständige Entsäuerung und Konservierung Ihrer Schätze. Gerne bieten wir Ihnen eine unverbindliche Vorführung an. Sprechen Sie mit uns!



Acidified paper will eventually destroy valuable books.

The unique, patented Datukom system ensures complete de-acidification and conservation. We are happy to offer a free demonstration. Just contact us for further information.

DATUKOM
Buchentsäuerung GmbH

Brettergartenstrasse 16
D-90427 Nürnberg
Telefon 09 11/3 26 30 10
und 01 70/3 47 01 40
Telefax 09 11/3 26 30 68

Suche über verschiedene Bibliothekskataloge dublette Einträge zu erwarten. Aus Sicht des Nutzers wünschenswert ist die Zusammenführung der Ergebnisse, in der Dubletten erkannt werden. Innerhalb des Z39.50-Standards ist eine Dublettenkontrolle bislang nicht vorgesehen. Diese wird erst in Ansätzen von der Entwicklergemeinde diskutiert.[16] Findet eine Dublettenkontrolle statt, so erfolgt sie auf Seiten des Clients, sei es eines speziellen Z39.50-Clients wie Liman Scout oder eines WWW-Z39.50-Gateways, außerhalb des Protokolls.

In dem EU-Projekt Universe, einer großen Z39.50 Anwendung, hat man sich die Zusammenführung von Ergebnissen aus heterogenen bibliographischen Datenbanken zum Ziel gesetzt. Das Projekt wurde im Juni 1999 beendet, ein umfassender Abschlußbericht für die Öffentlichkeit liegt jedoch noch nicht vor.[17] Ein anderes Z39.50 basiertes Informationssystem für den Bereich mathematischer Ressourcen, EULER, löst das Dublettenproblem über einen speziell für EULER konzipierten Matchkey. Dieser „Schlüssel" wird lokal in den Datensätzen gespeichert.[18] Diese Lösung ist für den KOBV schon deshalb nicht gangbar, da hier in die Datenstruktur und Indexierungspraxis der lokalen Systeme eingegriffen wird.
Die Dublettenprüfung kann nur auf der Suchmaschine erfolgen, die einerseits für Benutzer ein WWW-Z39.50-Gateway anbietet und gleichzeitig Server-Funktion für die Bibliothekare erfüllt.

Die Frage nach dem Recherchekomfort im Sinne von Registersuchen und Dublettenkontrolle muß nicht zuletzt unter dem Gesichtspunkt der geforderten Antwortzeiten betrachtet werden.

7. 5. Systemarchitektur: Quick Search in der Verteilten Suche

In dem Balanceakt zwischen der Offenheit auf der einen und Performanz als auch Skalierbarkeit auf der anderen Seite wurde nach intensiven, zum Teil kontroversen Diskussionen für die Suchmaschine eine Architekturentscheidung getroffen, die drei Arten von Suchen unterscheidet:

- IQuick Search über einen zentralen Gemeinsamen Index
- IParallele verteilte Suche über WWW für den Nutzer (WWW-Z39.50-Gateway)
- ISequentiell verteilte Suche über Z39.50 für den Bibliothekar

Um eine schnelle Antwortzeit von 3-5 Sekunden bei gleichzeitig hohem Recherchekomfort (Register / Dublettenkontrolle) zu garantieren, wird für einige der Zielsysteme eine Datenbank mit Indexinformationen zentral aufgebaut. Dieser Gemeinsame Index speist sich aus Daten lokaler Systeme, die damit nicht lokal verteilt, sondern zentral abgesucht werden. Als Ergebnis werden „Links" auf die Originalkatalogisate in den lokalen Systemen gesetzt. Dadurch wird für einen Teil der Bestände eine schnelle Suche (Quick-Search) garantiert. In diesem Zusammenspiel zwischen zentralen und virtuellen Komponenten sind zahlreiche Varianten denkbar. Unabhängig von dem lokal genutzten Bibliotheksinformationssystem, ist zu entscheiden, welche Bibliotheken ihre Daten in den Gemeinsamen Index einspielen. Vorstellbar wäre, die Daten der am häufigsten ausgewählten Zielsysteme in den Gemeinsamen Index zusammenzuspielen oder diejenigen aus Systemen mit der längsten Antwortzeit oder auch der lokal schwächsten Hardwareausstattung und Netzanbindung. Diesbezügliche Leitlinien wird der KOBV für den Verbund nach den ersten empirischen Erfahrungen und Nutzungsanalysen noch zu entwickeln haben.

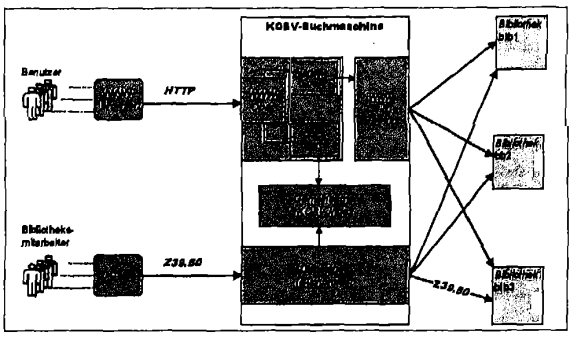


Abbildung 1: Architektur der Suchmaschine

Als grundsätzliche Philosophie des Verbundes wird an der verteilten Suche festgehalten. Die Kommunikation für Nutzer und Bibliothekare fußt dabei auf zwei unterschiedlichen Protokollen. Je nach Zugang, WWW für den Nutzer, Z39.50 für den Bibliothekar, unterscheidet sich die Recherche. Über das WWW-Z39.50-Gateway werden Anfragen parallel an die Zielsysteme weitergereicht (siehe oberer Teil in der Abbildung). Die Anfrage, die die Suchmaschine direkt über Z39.50 aus den lokalen Bibliothekssystemen erreicht, kann auch an mehrere Zielsysteme gerichtet sein. Die Suchmaschine übernimmt dabei die Verteilung der Anfragen an die angegebenen Zielsysteme (siehe unterer Teil in der Abbildung). Die Verteilung erfolgt sequentiell. Dabei entscheidet der Bibliothekar mit Hilfe eines (individuellen) Profils über die Reihenfolge, in der die Zielsysteme angefragt werden.
Für die verteilte Suche im WWW ist eine Dublettenkontrolle und das Mischen von Registern vorgesehen. Beides gilt allerdings zunächst nur für kleine Treffermengen. Eine Dublettenkontrolle ist für die sequentielle Suche obsolet geworden, da hier nur eine Ergebnismenge aus einem Zielsystem erwartet wird.

Nach diesem groben Überblick über die Architektur werden im folgenden die einzelnen Komponenten der KOBV-Suchmaschine noch einmal im Detail vorgestellt.

8. Parallele Verteilte Suche: Zugriff mit WWW-Browsern

Die KOBV-Suchmaschine fungiert als WWW-Z39.50-Gateway, wenn sie dem Nutzer die parallele Suche über mehrere Bibliothekskataloge anbietet. Ein „Katalog" kann dabei der Gemeinsame Index sein (siehe dazu Kapitel 10). Der Nutzer sucht über eine WWW-Oberfläche unabhängig vom Benutzerinterface und den Anfragesprachen der ausgewählten Datenbanken.
Ein WWW-Z39.50-Gateway war für das Aleph500-System der Version 12.1 bereits realisiert. Implementiert sind mit den Funktionalitäten (= Services) INIT, SEARCH und PRESENT die Basisfunktionen des Z39.50-Protokolls. Von den Zielsystemen wird als Ausgabeformat MAB2 erwartet. Ein MAB2-Satz aus einem Zielsystem wird temporär in das Aleph-Intern-Format umgewandelt und dem Nutzer wie ein Treffer aus einer Aleph-Bibliothek dargestellt.

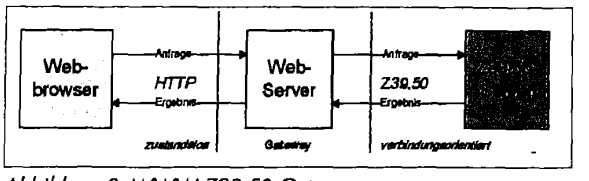


Abbildung 2: WWW-Z39.50-Gateway

Gesucht wird in lokalen Datenbanken. Die Pflege dieser Daten liegt in den Händen der lokalen Einrichtungen. Darin liegt der Vorteil und gleichzeitig die grundsätzliche Problematik einer verteilten Suche.
Die Heterogenität der Datenbanken geht einher mit unterschiedlichen Indexierungpraktiken und einer unterschiedlichen Anfragemächtigkeit der Systeme. Die Suchmaschine steht damit vor der Frage, welche Recherchemöglichkeiten sie anbieten kann. Schließlich steht und fällt die Anfragemächtigkeit des Gateways mit den Möglichkeiten der angesprochenen Zielsysteme. Die Suchmaschine kann also nicht „besser" sein als ihre Zielsysteme. Das gilt für die angebotenen Suchaspekte selbst wie für Trunkierungsmöglichkeiten und den Einsatz von Booleschen Operatoren.

Es ist vermutlich schwierig, dem Nutzer zu verdeutlichen, daß gewisse Suchaspekte von bestimmten Datenbanken nicht unterstützt werden, beispielsweise in Datenbank A mit der ISBN-Nummer gesucht werden kann, Datenbank B diesen Suchaspekt aber nicht kennt. Dasselbe gilt für die Unterstützung Boolescher Operatoren; so führt der Boolesche Operator AND-NOT in einer Datenbank zu einer Einschränkung der Treffermenge, eine andere ignoriert ihn jedoch.[19]

In diesem Zusammenhang kommt als besondere Schwierigkeit für den deutschsprachigen Raum hinzu, daß eine quasiverbindliche Umsetzung des nationalen Formates MAB auf den im Z39.50-Protokoll definierten Satz von Suchattributen (BIB1-USE-Attributes) bislang nicht existiert.[20] Das hat zur Folge, daß in manchen Datenbanken mit dem BIB1-USE-Attribut Personal Name (USE Attribut Type 1=1) nur Autoren gefunden werden, in anderen mit demselben Attribut zusätzlich auch über Personennamenschlagwörter gesucht wird.

Vor dem Hintergrund dieser grundsätzlichen Problematik hat die KOBV-Projektgruppe in der ersten Version der Suchmaschine einen pragmatischen Weg eingeschlagen, denjenigen des „kleinsten gemeinsamen Nenners". Nur diejenigen Recherchen werden seitens des Gateways angeboten, die von allen angebotenen Zielsystemen unterstützt werden. In dieser ersten Version werden die Anfragen in der Regel unabhängig von den Besonderheiten des jeweiligen Zielsystem in die Z39.50-Anfragesprache umgesetzt. In der nächsten Version der Suchmaschine ist die Möglichkeit gegeben, die Anfragen spezifischer - d.h. abgestimmt auf die ausgewählte Datenbank - zu transformieren.

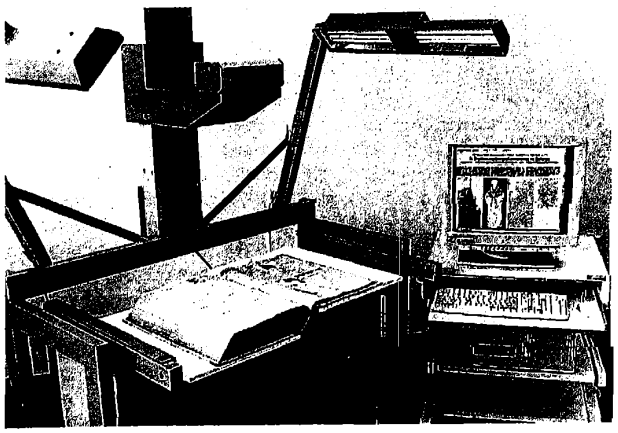
Eine Verbesserung der Rechechequalität kann mittelfristig durch ein für KOBV-Bibliotheken verbindliches Z39.50-Profil erreicht werden. Dieses Profil legt neben der Umsetzung von MAB nach BIB1 auch Defaulteinstellungen (z.B. Rechtstrunkierung ja / nein) fest. Ein solches Profil wäre für die Z30.50-Kommunikation auf nationaler Ebene wünschenswert.

8.1. Die WWW-Oberfläche der KOBV-Suchmaschine

Die WWW-Oberfläche für die KOBV-Suchmaschine entspricht in ihren Grundfunktionalitäten dem WWW-OPAC des Aleph 500-Systems. Dieser OPAC gibt eine dreiteilige Framestruktur

Neuer Aufsicht-Scanner/Kopierer für Bücher und Zeitungen

Ein neuartiges Scan-System OMNISCAN 5000 zum Digitalisieren und Kopieren, z. B. von Büchern und Zeitungen.



- Bequemes Scannen des Buches ohne mühsames Wenden – **Seiten nach oben!** – und ohne Verzerrung.
- Scannt beide Seiten gleichzeitig.
- Vorlagenformat max. DIN A1 (841 x 594 mm).
- Bis zu 256 Graustufen
- Schnelles Scannen: ca. 3 Sekunden (ohne Prescan)
- Auflösung: CCD-Element mit 5000 Pixel.
- Farbscanner, Format A0 auf Anfrage.

ZEUTSCHEL GmbH
Heerweg 2, 72070 Tübingen

ZEUTSCHEL

E-Mail: zeutschel@t-online.de
Internet: www.zeutschel.de

vor. Der Kopf enthält eine Leiste mit dem Hauptmenü, mit dessen Hilfe der Benutzer zwischen den verschiedenen Anwendungen wechseln kann. Im rechten und gleichzeitig größten Frame befindet sich die Arbeitsfläche des Benutzers mit der Eingabemaske und der Ergebnis-Anzeige. Im linken Frame mit dem Funktionsmenü kann der Benutzer auswählen, welche Funktionen er ausführen möchte. Dieser vorgegebene Rahmen, der die genuine Funktionalität des Aleph500-WWW-OPACs beinhaltet, mußte bei der Gestaltung der Suchmaschinen-Oberfläche beibehalten werden. Dagegen wurden die vielfältigen Optionen, die der Aleph500-WWW-OPAC in seiner Grundausführung bietet, stark reduziert, um eine möglichst klar strukturierte Oberfläche und einfach handhabbare Funktionen zu bieten, die sich auch dem ungeübten Nutzer auf den ersten Blick erschließen. Aus dem gleichen Grund ist die Begrifflichkeit bewußt einfach gehalten, und es wurde so weit wie möglich auf Fachtermini (sowohl bibliothekarische als auch EDV-technische) verzichtet.

Die nächste Version der Suchmaschine wird einen deutlich höheren Recherchekomfort bieten als die erste Version, die hier vorgestellt wird. Dabei ist neben der Suche über alphabetische Registereinträge, die für einen definierten Suchraum dynamisch zu Anfragezeit aufgebaut werden, in erster Linie die Zusammenführung von Ergebnissen aus den verschiedenen Datenbanken geplant. Ziel ist es, eine Dublettenkontrolle durchzuführen, analog zu dem Verfahren, das beim Aufbau des Gemeinsamen Indexes zum Einsatz kommt (siehe dazu Kapitel 10.2).

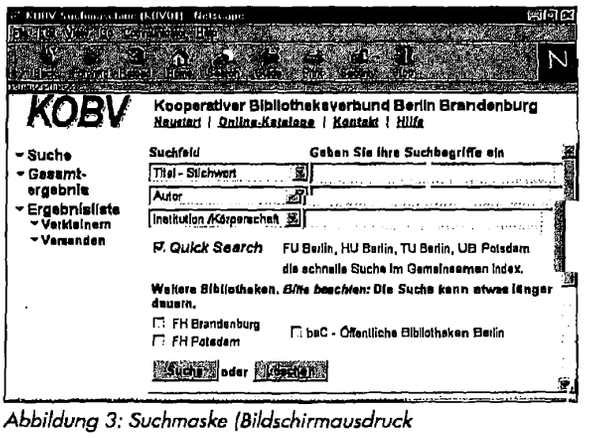


Abbildung 3: Suchmaske (Bildschirmausdruck

Die Suchmaske der KOBV-Suchmaschine ist in mehrere Bereiche aufgeteilt: im obersten befindet sich das Eingabeformular, darunter die Liste der auswählbaren Bibliotheken und der Sende-Button.
Das Eingabeformular für die Suche besteht aus drei Feldern. Innerhalb eines Feldes ist die Eingabe mehrerer Wörter zulässig, die mit den Booleschen Operatoren AND oder OR verknüpft werden können. Einzelwörter können rechts trunkiert werden. Untereinander stehen die drei Felder durch den Booleschen Operator AND in Verbindung. Listboxen links von der Eingabezeile enthalten die suchbaren Kategorien. Gesucht werden kann mit Titelstichwörtern nach Sachtiteln, Serien und Zeitschriften, mit Autoren- und Körperschaftsnamen sowie mit Schlagwörtern.[21]
Im mittleren Bereich des Fensters kann der Nutzer auswählen, in welchen Zielsystemen er suchen möchte. Der Quick-Search und die Bibliotheken können sowohl einzeln angewählt werden als auch zusammen in beliebigen Kombinationen. Voreingestellt ist die Suche über den Quick-Search, der schnellen Suche im Gemeinsamen Index. Die Besonderheit dieser Funktion besteht darin, daß nur auf einen Server zugegriffen wird, nicht auf mehrere wie bei der verteilten Suche, wodurch eine effizienter Zugriff garantiert wird. Um diese Besonderheit hervorzuheben, ist der Quick-Search von den übrigen Katalogen etwas abgesetzt. In der ersten Version der KOBV-Suchmaschine, die Anfang November 1999 freigegeben wird, ist die Anzahl der angesprochenen Datenbanken noch recht beschränkt. Die Zahl der Bibliotheken wird sich beim Ausbau der KOBV-Suchmaschine bis Ende des Jahres und im nächsten Jahr stark erweitern.[22]

Das Gesamtergebnis seiner Suche wird dem Nutzer im rechten Frame angezeigt. Die Ergebnisse aus den verschiedenen ausgewählten Katalogen werden bislang noch in einzelnen Sets angezeigt. Eine Zusammenführung der Ergebnisse aus unterschiedlichen Datenbanken erfolgt in dieser Version der KOBV-Suchmaschine nicht.

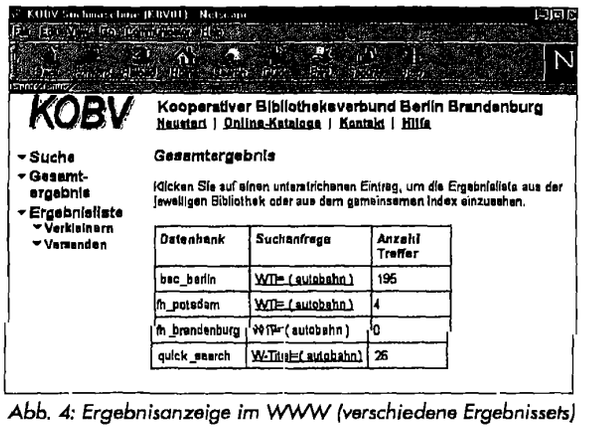


Abb. 4: Ergebnisanzeige im WWW (verschiedene Ergebnissets)

Ein Klick auf die Liste eines Kataloges im Gesamtergebnis führt zur Ergebnisliste einer Kurztitelliste der ausgewählten Bibliothek. Ist diese Ergebnisliste zu groß, kann sie über weitere (selektive) Attribute wie Titelstichwort, Verlag, Erscheinungsort und Erscheinungsjahr eingeschränkt werden. Trefferlisten können in den unterschiedlichen Formaten lokal abgespeichert und per E-Mail verschickt werden.
In der Ergebnisliste kann der Nutzer einen Titel auswählen und gelangt dann zur vollen bibliographischen Information. Er kann sich einen Titel in verschiedenen Formaten anzeigen lassen: neben dem Standardformat das Katalogkarten- oder ISBD-Format und - als Referenz an die bibliothekarischen Nutzer - ein Feldnummern-Format mit den MAB2-Kategorien. Im Standardformat hat der Nutzer die Möglichkeit, über einen sogenannten WEB-LINK in den WWW-OPAC eines lokalen Systems zu „springen". Dort kann er alle Funktionalitäten nutzen, die ihm dieses lokale System bietet - bis hin zur Initiierung einer Bestellung.

Mit dem WEB-LINK und der Verzweigung in die angeschlossenen lokalen Systeme ist in der KOBV-Suchmaschine eine Funktion realisiert, die bislang keine andere Verbundsoftware bietet.

8.2 Verzweigen in die lokalen OPACS (WEB-LINK)

Die Suchmaschine erhält von den Zielsystemen über Z39.50 bibliographische Sätze, die zunächst keine Information über die jeweiligen Exemplare, geschweige denn über deren aktuellen Ausleihstatus, enthalten. Eben daran aber ist der Nutzer interessiert, sucht er doch in der Regel nicht nach bibliographischen Einträgen, sondern nach Dokumenten, die er bestellen kann.

Die besondere Qualität der KOBV-Suchmaschine liegt nun darin, daß sie den Nutzer direkt in die lokalen Opacs derjenigen Bibliotheken, die den gefundenen Titel besitzen, weiterleitet.

Das Verzweigen in die lokalen Opacs erfolgt über einen WEB-Link. Dieser Link führt den Nutzer zu genau dem Titel dem lokalen WEB-Opac, den er über die Suchmaschine gefunden hat. Im lokalen System kann der Nutzer dann das gewünschte Exemplar bestellen bzw. für die Bestellung vormerken lassen. Unter der Voraussetzung natürlich, daß er über die entsprechende Berechtigung verfügt.

Dieser WEB-Link basiert auf einer sogenannten SID, einer für den Bereich des KOBV eindeutigen Identnummer. Die SID gliedert sich in drei Unterfelder. Das erste Unterfeld enthält die Information, ob es sich um einen Link zu einem Aleph-System oder zu einem anderen System handelt. Diese Information wird für die interne Verarbeitung benötigt. Das zweite Unterfeld enthält ein eindeutiges Bibliothekskürzel, das zentral festgelegt wird. Dieses Kürzel wird in den vollständigen Namen der Bibliothek umgesetzt, die der Benutzer im WWW zu sehen bekommt. Im dritten Unterfeld wird eine Nummer aus dem lokalen System gespeichert, die den Titelsatz in dem jeweiligen lokalen System eindeutig identifiziert. Zu jedem Titelsatz, den die Suchmaschine findet, erzeugt sie eine solche eindeutige SID.

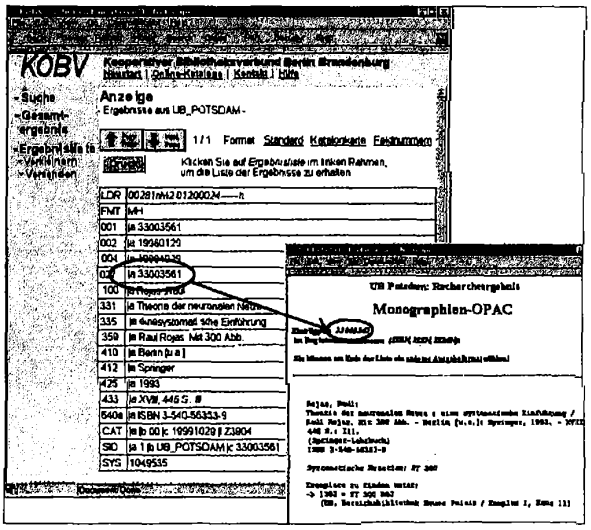


Abbildung 5: WEB-LINK-Funktion (Bildschirmausdruck)

Mit dem Klicken auf einen WEB-Link wird ein spezielles, für die einzelne Bibliothek angepaßtes CGI-Scipt gestartet, das für den Benutzer die Suche nach dem erzielten Treffer im lokalen OPAC übernimmt. In diesem CGI-Script ist vordefiniert, in welchem Index die in der SID gespeicherte lokale Identnummer abgelegt und für die KOBV-Suchmaschine recherchierbar ist.
Noch können nicht in allen am KOBV beteiligten Bibliotheken Bestellungen und Vormerkungen über das WWW abgesetzt werden. Neben technischen Schwierigkeiten stehen zusätzlich auch organisatorische (z.B. auch unterschiedliche Benutzerausweise) so manchem Benutzerwunsch nach einer schnellen Lieferung des gefundenen Dokuments entgegen. Auf der zentralen technischen Ebene allerdings ist mit dem WEB-LINK auf einfache Weise bereits jetzt ein Grundstein gelegt worden, um mit dem Kooperativen Bibliotheksverbund Berlin-Brandenburg einen Ausleihverbund zu realisieren.

9. Sequentielle verteilte Suche: Zugriff über Z39.50-Clients

Ausführlich wurde diskutiert, inwieweit analog zur Suche über die WWW-Schnittstelle die parallele Suche mit anschließender Dublettenkontrolle auch für die von Bibliothekaren genutzte Z39.50-Client-Schnittstelle realisiert werden kann.[23] Das würde bedeuten, den Z39.50-Standard, in dem nur die Ausgabe eines Ergebnissets von einem Zielsystem (= Target) vorgesehen ist (Punkt-zu-Punkt bzw. 1-zu-1-Verbindung), neu zu interpretieren und die Zusammenführung von mehreren Ergebnissen im Aleph-System erfolgen zu lassen (1-zu-N-Verbindungen).
In dieser reinen Z39.50 zu Z39.50 Kommunikation agiert der Z-Server der Suchmaschine für die Z-Clients als Dispatcher, der die Anfragen an die Zielsysteme weiterleitet.

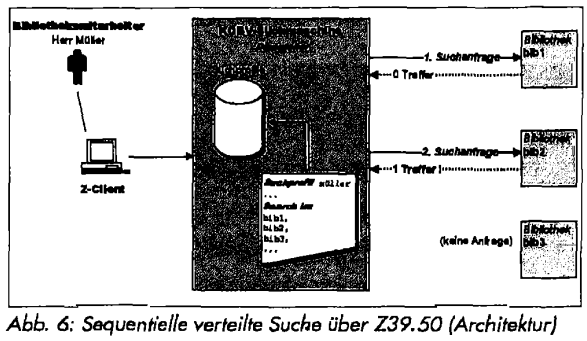


Abb. 6: Sequentielle verteilte Suche über Z39.50 (Architektur)

Die Suche erfolgt anhand eines Suchprofils. Der Benutzer legt in einem Suchprofil fest, in welchen Bibliotheken er recherchieren möchte. Dabei definiert er die Reihenfolge, in der die Anfragen abgearbeitet werden. Wird in einer Bibliothek ein Treffer gefunden, so zeigt die Suchmaschine die dazugehörige Ergebnisliste an und verzichtet auf die Suche in weiteren Bibliotheken. Damit bekommt der Nutzer jeweils nur eine Treffermenge aus einem Zielsystem, eine Dublettenprüfung sowie das Sortieren von Ergebnismengen aus unterschiedlichen Zielsystemen ist in dieser Konzeption obsolet geworden.
Möchte der Benutzer auch Treffer aus anderen Bibliotheken sehen, kann er sein Profil ändern bzw. mit einem zweiten Profil arbeiten, in dem die Reihenfolge der Zielsysteme eine andere ist. Die Suchprofile werden individuell definiert, über eine WEB-Schnittstelle erstellt und gepflegt und zentral auf der Suchmaschine gespeichert.

Die Recherchemöglichkeiten werden von den jeweiligen Zielsystemen vorgegeben. Sie werden im Einzelfall größer sein, als die über die WWW-Schnittstelle der KOBV-Suchmaschine angebotenen. So ist über die Z-Client-Schnittstelle die Suche über Registereinträge bereits jetzt möglich, da die angesprochenen Z-Server die Funktionalität SCAN durchweg implementiert haben. Neben den unterschiedlichen Stärken auf seiten der angesprochenen Server stehen auch unterschiedliche Möglichkeiten auf seiten der einzelnen Clients. Eine Hilfestellung für den Benutzer können in diesem Zusammenhang ausführliche Dokumentationen bieten.[24]

Innerhalb der Z39.50 zu Z39.50 Kommunikation stellt der Gemeinsame Index als Zielsystem eine Besonderheit dar, da hier nicht der vollständige Datensatz vorhanden ist (siehe dazu das folgende Kapitel 10.1). Die sequentiell verteilte Suche befand sich im Oktober 1999 - als dieser Artikel geschrieben wurde - noch im Test, so daß Erfahrungen aus der Praxis noch nicht vorliegen.

10. Gemeinsamer Index (Quick Search)

Der sogenannte Gemeinsame Index verfolgt das Ziel, für einen gewissen Datenbestand eine performante Suche zu garantieren. Daten aus unterschiedlichen Quellen werden zusammengespielt und dabei auf dublette Einträge überprüft.
Aus Sicht der KOBV-Suchmaschine ist der zentrale Gemeinsame Index ein weiteres lokales Zielsystem, das in der Suche gleichberechtigt neben den anderen lokalen Systemen steht. Es handelt sich in gewisser Weise um einen „Rumpfkatalog", der nach der Typologie aus den eingangs zitierten Empfehlungen zwischen Modell 4 und 5 anzusiedeln wäre. In der ersten Version dieses Index wurden die Datensätze der Bibliothekssysteme der Humboldt-Universität (HU), der Freien Universität (FU) der Technischen Universität (TU) und der Universität Potsdam eingespielt. Für den Aufbau des Index werden die Lade- und Indexierungsprozeduren genutzt, die die Software Aleph500 bietet.

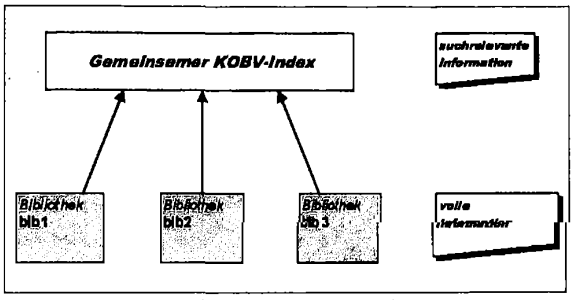


Abbildung 7: Architektur Gemeinsamer Index

10.1 Indexinformationen - Datenstruktur

Im Gemeinsamen Index wird nicht der vollständige Titelsatz zentral gehalten und indexiert, sondern nur bestimmte Elemente. Für den formalbibliographischen Teil des Datensatzes sind es in erster Linie die suchrelevanten Elemente. Daneben sind es diejenigen Felder, die für die Dublettenkontrolle und die Standardanzeige benötigt werden. Zur Zeit wird mit der Anzahl der zentral indexierten, formalbibliographischen Felder noch experimentiert.
Wird ein Datensatz als dublett erkannt, werden nur die oben genannten formalbibliographischen Elemente eines Satzes indexiert (zur Dublettenerkennung siehe Kapitel 10.2). Welcher Satz dabei der bevorzugte ist, legt ein parametrisierbarer Merge-Algorithmus fest. Anders verhält es sich mit Schlagwörtern (hier nur RSWK-Schlagwörter). Um eine Suche nach möglichst vielen Schlagwörtern zu ermöglichen, werden diese immer gespeichert. Erkennt die Suchmaschine zwei Datensätze als dublett, die beide Schlagworte enthalten, enthält ein Datensatz im Gemeinsamen Index die Schlagwörter aus beiden Sätzen.
Ein Eintrag im Gemeinsamen Index enthält neben formalen Elementen und Sacherschließungselementen auch einen Schlüssel (SID), der den Titelsatz in einer oder mehreren lokalen Bibliotheken eindeutig identifiziert

Bei der Suche über das WWW wird dem Benutzer der reduzierte Datensatz mit den entsprechenden Links in die lokalen Opacs angezeigt. Anders ist es bei der Recherche über Z39.50 für den Bibliothekar. Da der anfragende Katalogisierer die vollständige Information benötigt, wird der entsprechende Datensatz (über die SID) aus dem jeweiligen lokalen System geholt.

BIBLIOTHEKS-BUCHBINDEREI

BUCH-RESTAURIERUNG

BUCHBINDEREI
OBERMEIER

Inh. Werner Obermeier
Buchbindermeister, Diplomierter Buchrestaurator
Benzstr. 2 · 84056 Rottenburg / Laaber
Telefon 08781/5 50 · Fax 08781/25 32
Niederlassung München: Heßstrasse 14 · 80799 München
(ehem. Handbuchbinderei Fritz Schmidt)
Telefon 089/2800116 ·Fax 089/28659049

10.2 Dublettenbehandlung

Die Dublettenkontrolle basiert auf einer Spezifikation der KOBV-Projektgruppe und wurde von Ex Libris vollständig neu entwickelt.[25]
Um nicht jeden Satz mit jedem vergleichen zu müssen (1 zu N Vergleich), sucht das Programm im ersten Schritt nach potentiellen Dubletten. Dazu wird zunächst mit einem Wort aus dem Titel nach potentiellen Kandidaten gesucht. Ist die Treffermenge größer als 50, erzeugt das System eine neue Anfrage mit einem weiteren Wort aus dem Titel. Liegt nach der Suche mit sämtlichen Wörtern aus dem Titel die Anzahl der Treffer immer noch über 50, werden Wörter aus dem Autorenfeld respektive aus dem Körperschaftsfeld hinzugezogen, um die Treffermenge weiter einzuschränken.[26]

Ein Datensatz wird nach der derzeitigen Parametrisierung maximal mit 100 Datensätzen verglichen.
Die Dublettenprüfung der ersten Version des Gemeinsamen Index zieht die folgenden (MAB) Felder heran:[27]

100: Verfasser
200: Körperschaft
331: Hauptsachtitel
335: Zusatz zum Hauptsachtitel
403: Auflagenbezeichnung (nur numerische Bestandteile)
410: Verlag (bis zu 5 Zeichen)
412: Verlagsort (bis zu 5 Zeichen)
425: Erscheinungsort (nur numerische Bestandteile)
433: Umfangsangabe (nur numerische Bestandteile)
455: Bandangabe (nur numerische Bestandteile)
540: ISBN

Es hat sich gezeigt, daß die Dublettenerkennung insbesondere bei Unterreihen noch mangelhaft ist. Bei einem Neuaufbau muß zusätzlich der Ansetzungstitel herangezogen werden.

Die oben aufgelisteten Felder werden nur berücksichtigt, wenn sie in beiden Datensätzen vorhanden sind. Eine Ausnahme bildet die Auflagenbezeichnung. Ist die Auflagenbezeichnung in einem Datensatz nicht vergeben, wird sie dem Satz mit dem Wert 1 systemseitig hinzugefügt.

Der Vergleich selbst arbeitet exakt (auf Gleichheit) oder tolerant (auf Ähnlichkeit). Exakt werden diejenigen Felder verglichen, in denen ein numerischer Wert erwartet wird (z.B. ISBN, Erscheinungsjahr). Tolerant verglichen werden Textfelder wie Verfasser, Körperschaft und Titelangaben. Ziel dieses „toleranten Vergleichs", bei dem ein sogenannter parametrisierbarer Trigram-Algorithmus zum Tragen kommt, ist es, kleinere Unterschiede in den Zeichenstrings wie Tippfehler und unterschiedliche Abkürzungen zuzulassen.

Jedem Feld ist ein positiver und ein negativer Wert zugeordnet. Ein positiver Wert spricht bei Gleichheit für eine Dublette, ein negativer bei Ungleichheit gegen eine Dublette. Ein positiver wie ein negativer Schwellwert ist definiert worden. Damit gilt nur als dublett, wenn der erreichte Gesamtwert des Datensatzes nach dem Vergleich gleichzeitig über dem positiven und unter dem negativen Grenzwert liegt.

Unsere bisherigen Erfahrungen haben gezeigt, daß sich mit diesem aufwendigen Verfahren prinzipiell qualitativ gute Ergebnisse erzielen lassen. In der ersten Version lag die Fehlerrate dennoch bei überschlägig 2%. Das heißt, 2% derjenigen Datensätze, die als dublett erkannt wurden, hätten nicht zusammengeführt werden dürfen. Die anschließende Analyse der Fehlerfälle hat jedoch ergeben, daß sich der Großteil dieser Fehler durch Parameteränderungen insbesondere bei der Festlegung der Gewichte korrigieren läßt.

Hinsichtlich der Performanz ist deutlich geworden, daß die Suche nach Kandidaten zeitaufwendiger ist, als zunächst angenommen. Die Optimierungen des Verfahrens werden an dieser Stelle ansetzen müssen.

10.3 Update

Da die Entscheidung, welche Bibliothek ihre Daten in den Gemeinsamen Index hochspielt, nicht von dem lokal eingesetzten Bibliothekssystem abhängig ist, wurde zusätzlich zu dem vorhandenen Aleph-spezifischen Replikationsverfahren ein asynchrones Updateverfahren auf der Basis des Z39.50 - Protokoll entwickelt.

Das Updateverfahren arbeitet in zwei Schritten. Zunächst signalisiert ein lokales System der Suchmaschine jeden neuen, korrigierten oder gelöschten Datensatz. Ein solches Signal bestehend aus

- einer kodierten Z39.50-Suchanfrage,[28] die diesen Datensatz eindeutig kennzeichnet (i.a. ist dies die Systemnummer)
- einem Aktionscode (NEW, UPD, DEL)
- dem vom KOBV vergebenen Bibliothekskennzeichen sowie
- einem Datum- und Zeitstempel.

Diese Signale werden beispielsweise täglich oder auch wöchentlich gesammelt und per FTP an die Suchmaschine übermittelt.

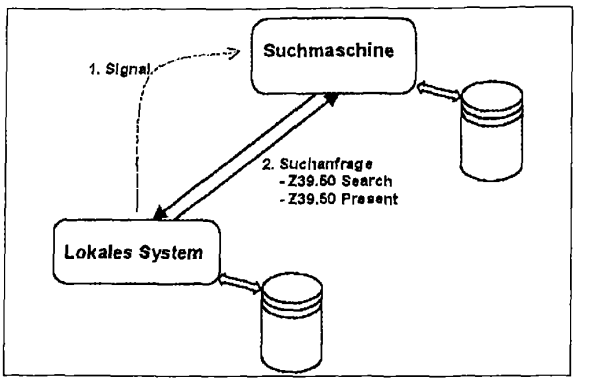


Abbildung 8: Update des Gemeinsamen Indexes über Z39.50

Im zweiten Schritt generiert die Suchmaschine aus Bibliothekskennzeichen und kodierter Suchanfrage die SID, die den Datensatz KOBV-weit eindeutig kennzeichnet. Beim Einfügen (Aktionscode NEW) bzw. Update (UPD) führt die Suchmaschine die im Signal kodierte Suchanfrage (Z39.50 Search) aus und erhält den Datensatz (Z39.50 Present) vom lokalen System geliefert. Unter Berücksichtigung des oben beschriebenen Dublettenerkennungsverfahrens wird der Datensatz in den Gemeinsamen Index eingefügt bzw. korrigiert. Dabei erhält ein dubletter Satz einen weiteren Link mit seiner SID. Beim Löschen (DEL) wird der entsprechende Link aus dem Gemeinsamen Index entfernt.

11. Ausblick

Die Aufgaben, die sich die Projektgruppe am ZIB bis zum Ende der Projektlaufzeit Ende 2000 stellen muß, lassen sich mit den Stichworten Konsolidierung, Optimierung und Weiterentwicklung der KOBV-Suchmaschine beschreiben.

Vorgesehen ist die schrittweise Integration weiterer Bibliotheken aus der Region. Dabei stellt die Integration der Staatsbibliothek zu Berlin aufgrund des Umfangs ihrer Bestände eine besondere Herausforderung dar. Technisch offen ist die Architektur für Bibliotheken jeder Sparte und Größe. Auch Subverbünde, wie beispielsweise der im Aufbau befindliche Verbund Öffentlicher Bibliotheken Berlins (VÖBB), können integriert werden. Das gleiche gilt für Anbieter von Fachinformationen, soweit deren Systeme über standardisierte Schnittstellen verfügen.[29] Ziel des Kooperativen Bibliotheksverbunds Berlin Brandenburg ist es, sich zu einem möglichst umfassenden Nachweisinstrument für die Region zu entwickeln.

Die vorgesehenen Optimierungen und Weiterentwicklungen der Suchmaschine gehen Hand in Hand. Der Recherchekomfort ist deutlich zu verbessern. So soll eine Dublettenprüfung als Option für den Benutzer realisiert werden. Angestrebt wird darüber hinaus die Suche über Registereinträge. Für eine vordefinierte Anzahl von Bibliotheken sollen dabei die Registereinträge dynamisch zusammengeführt werden.

Die Entwicklung der KOBV-Suchmaschine wird nach Projektplan im Frühjahr 2000 abgeschlossen sein. Ein Schwerpunkt liegt dabei in der funktionalen Integration eines Anbieterverzeichnisses.[30] Ziel ist es, virtuelle Fachbibliotheken wie zum Beispiel eine virtuelle Fachbibliothek der Mathematik zu ermöglichen. Dazu soll der Suche nach einzelnen Dokumenten eine Suche nach Anbietern von Dokumenten (≈ Bibliotheken) vorgeschaltet werden. Indexierte Metadaten über die jeweiligen Anbieter - formale Beschreibungen wie der Name der Bibliothek und sachliche wie

die Klassifikation der Bibliotheksbestände nach einem Fachgebietsverzeichnis - sollen den Benutzer bei seiner Auswahl von geeigneten Zielsystemen unterstützen. Eine Übernahme der Ergebnisse in individuelle Suchprofile ist vorgesehen.

Organisatorisch steht die Überführung des Kooperativen Bibliotheksverbunds Berlin-Brandenburg in den Regelbetrieb bevor. An den Voraussetzungen für die Institutionalisierung des KOBV für den 1. Januar 2001 wird seitens der Träger im Moment gearbeitet.

Die Autoren gehen davon aus, daß das System kontinuierlich weiterentwickelt wird, um dem Nutzer adäquate Angebote unterbreiten zu können und Schritt zu halten mit der rasanten technologischen Entwicklung. Die Vorstellung, die zu Beginn des Projektes von den Vätern des KOBV, den Experten, im Interesse der Nutzer formuliert wurde, einen Paradigmenwechsel bei den Bibliotheken einzuläuten „von der klassischen Versorgungseinrichtung zum Informationsvermittler" ist noch längst nicht erreicht. Das Fundament indes ist gelegt.

Abkürzungsverzeichnis

AACR	Anglo-American Cataloguing Rules
BGK	Berliner Gesamtkatalog
BVB	Bibliotheks-Verbund Bayern
BVBB	Bibliotheksverbund Berlin-Brandenburg
DBI	Deutsches Bibliotheksinstitut
DFG	Deutsche Forschungsgemeinschaft
FTP	File Transfer Protocol
GBV	Gemeinsamer Bibliotheksverbund der Länder Bremen, Hamburg, Mecklenburg-Vorpommern, Niedersachsen, Sachsen-Anhalt, Schleswig-Holstein und Thüringen
GKD	Gemeinsame Körperschaftsdatei
KOBV	Kooperativer Bibliotheksverbund Berlin-Brandenburg
MAB	Maschinelles Austauschformat für Bibliotheken
PND	Personennamen Datei
SWB	Südwestdeutscher Bibliotheksverbund
SWD	Schlagwortnormdatei
US-MARC	Machine-Readable Cataloging
VÖBB	Verbund Öffentlicher Bibliotheken Berlins
WWW	World Wide Web
Z39.50	Standardisiertes Protokoll für den Bereich des Information Retrieval, ANSI/NISO-Standard
ZIB	Konrad-Zuse-Zentrum für Informationstechnik Berlin

Abbildungsverzeichnis

Abbildung 1: Architektur der Suchmaschine
Abbildung 2: WWW-Z39.50-Gateway
Abbildung 3: Suchmaske (Bildschirmausdruck)
Abbildung 4: Ergebnisanzeige im WWW (verschiedene Ergebnissets)
Abbildung 5: WEB-Link-Funktion
Abbildung 6: Sequentiell verteilte Suche über Z39.50
Abbildung 7: Architektur Gemeinsamer Index
Abbildung 8: Update des Gemeinsamen Indexes über Z39.50

[1] Dugall, Berndt; Hebgen, Michael; König, Wolfgang: Empfehlungen zur zukünftigen Struktur der Informationssysteme der wissenschaftlichen Bibliotheken des Landes Berlin unter Berücksichtigung der wissenschaftlichen Bibliotheken des Landes Brandenburg, Januar 1997. Elektronisch verfügbar unter *www.kobv.de/docs/*

[2] Eine Kritik an der Überbürokratisierung des deutschen bibliothekarischen Regelwerkes, das FremddatenÜbernahme aus dem anglo-amerikanischen Raum erschwere und den Nutzer außer Acht lasse, hatten Hebgen und König bereits früher an anderer Stelle geäußert. Die Autoren schlagen vor, das anglo-amerikanische Regelwerk AACR zu übernehmen. Ob damit allerdings der gewünschte Effekt erreicht werden kann, darf bezweifelt werden, da dieses Regelwerk mindestens ebenso umfangreich und kompliziert ist wie das deutsche. Vgl. dazu: Hebgen, Michael: Quo vadis Deutsches Bibliothekswesen? In: Wirtschaftsinformatik. - Braunschweig; Wiesbaden. - 38 (1996), 3, S. 337. König, Wolfgang: Bibliotheksinformationssysteme : behindern traditionelle Denkweisen moderne Konzepte? In: Wirtschaftsinformatik. - Braunschweig ; Wiesbaden. - 38 (1996), 3, S. 335 - 336

[3] Diedrichs vom GBV weist in diesem Zusammenhang nicht zu Unrecht darauf hin, daß mit einer personell bewußt klein gehaltenen Zentrale, mehr Aufgaben in den einzelnen lokalen Einrichtungen zu bewältigen sind.
Zu der Kontroverse über das Konzept des Kooperativen Bibliotheksverbundes Berlin-Brandenburg siehe auch die in der Zeitschrift Weitblick erschienenen Stellungnahmen. Weitblick: Mitteilungsblatt der Bibliotheken in Berlin und Brandenburg. 1997, 3, S. 31-35 und 1997,4, S. 32-44. Zu erreichen über:
http://www.kobv.de/project/disc/

[4] Die Ergebnisse dieser Arbeitgruppe sind nachzulesen bei: Lügger, Joachim (Red.): Kooperativer Bibliotheksverbund Berlin- Brandenburg: Lokale Konzepte und technische Schnittstellen, ZIB TR 97-14. - Mai 1997. Im Volltext unter: *http://www.kobv.de/docs/*

[5] Informationen über das Konrad-Zuse-Zentrum finden Sie im Internet unter *http://www.zib.de*

[6] Im Math-Net-Projekt wurde erfolgreich der Versuch unternommen, die verteilt auf universitären Rechnern in Deutschland liegenden, für die Mathematik relevanten Ressourcen - Publikationen, Software oder Homepages - über eine Suche recherchierbar zu machen. Dabei wird die Harvest-Technologie eingesetzt. Das System Harvest, bestehend aus den Komponenten Broker und Gatherer, kann auf der lokalen auf der überregionalen (zentralen) Ebene Indexe erzeugen. Die Gatherer sammeln die Informationen von WWW-, Hyper-G- und FTP-Servern lokal oder über das Internet. Die dadurch gebildeten Suchindexe werden von Brokern lokal oder auch überregional (zentral) zu umfassenderen Indexen zusammengefaßt. Der Benutzer sucht mit seinem WWW-Client (Browser) über das HTTP-Protokoll in den maschinell zusammengeführten Indexinformationen. Nähere Informationen zum Math-Net-Projekt finden Sie im Internet unter http://www.math-net.de

[7] Vereinbarung zur Durchführung des Projektes Kooperativer Bibliotheksverbund Berlin-Brandenburg. 11. Mai 1998. Als Volltext verfügbar unter *http://www.kobv.de/docs/*

[8] Konzept zur zukünftigen Informationsstruktur der Bibliotheksregion Berlin-Brandenburg. Gemeinsames Landeskonzept. August 1998. Als Volltext verfügbar unter *http://www.kobv.de/docs/*

[9] Information zu Ex Libris finden Sie unter *http://www.exl.de* und unter *http://www.aleph.co.il*

[10] Zur gemeinsamen Nutzung für die KOBV-Bibliotheken werden darüber hinaus zentral Fremd- und Normdaten (GKD, SWD, PND) bereitgehalten. Diese Daten werden

in einem Aleph-System bereitgestellt und sind über Z39.50 zugreifbar. Dieser Fremd- und Normddatenserver wird von der Suchmaschine analog zu einem lokalen System behandelt. Auf diese Datenbestände wird hier nicht näher eingegangen.

[11] Eine Einführung in Z39.50 leistet Bernd Hergeth in: Z39.50 in Bibliotheken und im World-Wide-Web, 1996. Als Volltext verfügbar unter *http://www.ub.uni-dortmund.de/Inetbib/v_hergeth.htm*
Einen guten Überblick bietet die Broschüre von William Moen: The ANSI/NISO Z39.50 Protocol: Information Retrieval in the Information Infrastructure. Als Volltext verfügbar unter *http://www.cni.org/pub/NISO/docs/Z39.50-brochure/50.brochure.toc.html*
Die offiziellen Seiten der Z39.50-Implementors Group, die Maintenance-Agency-Page, pflegt die Library of Congress: *http://lcweb.loc.gov/z3950/agency/agency.html*

[12] Eversberg, B: Z39.50-Schnittstelle für das System allegro. Universitätsbibliothek Braunschweig, 30. Dezember 1998. Als Volltext verfügbar unter *http://www.biblio.tubs.de/allegro/z3950/*

[13] Das DBV-OSI-Projekt (Deutscher Bibliothekenverbund - Open Systems Interconnection) wurde von 1993-1997 vom Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie und der Deutschen Forschungsgemeinschaft (DFG) finanziert. Entwickelt wurde u.a. eine Z39.50-Server-Schnittstelle für die Verbunddatenbanken des Gemeinsamen Bibliotheksverbundes (GBV), des Südwestverbundes (SWB) und des Bayerischen Bibliotheksverbundes (BVB). Siehe dazu auch: *http://www.ddb.de/partner/dbv-osi_ii.htm*

[14] Zum Zusammenspiel der Protokolle Z39.50 und HTTP siehe: Hammer, Sebastian und Favaro, John: Z39.50 and the World Wide Web In: D-Lib Magazine - The Magazine of Digital Library Research, March 1996. Im Volltext verfügbar unter: *http://mirrored.ukoln.ac.uk/lisjournals/dlib/dlib/dlib/march96/briefings/03indexdata.html*

[15] Siehe dazu auch Testergebnisse in Schneider, Wolfram: Ein verteiltes Bibliotheks-Informationssystem auf Basis des Z39.50 Protokolls. Diplomarbeit. TU-Berlin. 1999, S.20f. Als Preprint mit der Nummer SC99-21 im Volltext verfügbar unter: *http://www.zib.de/bib/pub/pw/index.de.html*

[16] Siehe: Z39.50-1995 Amendment 2: Z39.50 Duplicate Detection Service, April 1999 *http://lcweb.loc.gov/z3950/agency/amend/dedup.html*

[17] Nähere Informationen zum UNIVERSE-Projekt finden Sie im Internet unter *http://www.fdgroup.co.uk/research/universe/*

[18] Nähere Informationen zum EULER-Projekt finden Sie im Internet unter *http://harvest.sub.uni-goettingen.de/euler/project/euler.html*

[19] Die Problematik ist eindrücklich beschrieben in Lunau, Carrol D.: The Need for an International Z39.50 Profile for Searching Virtual Catalogues, Oktober 1998. Als Volltext verfügbar unter *http://www.nlc-bnc.ca/resource/vcuc/profiss.pdf*

[20] Für den anglo-amerikanischen Raum ist eine Art Quasi-Standard von der Library of Congress vorgelegt worden. In „Attribute Set BIB-1 (Z39.50-1995): Semantics" vom September 1995 finden sich als semantische Spezifizierung die Umsetzung des BIB1-Attribut-Sets auf US-MARC. Als Volltext verfügbar unter: *ftp://ftp.loc.gov/pub/z3950/defs/bib1.txt*

[21] Intensiv diskutiert wurde in der KOBV-Projektgruppe die Frage, ob man angesichts der Tatsache, daß nur ein Bruchteil der Bestände in den Bibliotheken verbal verschlagwortet ist, die Schlagwortsuche überhaupt anbieten soll. Da die Schlagwortsuche andererseits eine wertvolle Einstiegshilfe bietet, hat man sich letztendlich dafür entschieden, diese Option anzubieten und in den Hilfetexten auf die Problematik zu verweisen. Diese grundsätzlichen, ebenso in den lokalen OPACs virulenten Fragen, potenzieren sich auf der Ebene der Suchmaschine und sind mit technischen Mitteln alleine nicht zu lösen.

[22] Mit Zunahme der Zielsysteme sind individuelle Suchprofile für Benutzer vorgesehen. Für die nächste Version wird die Möglichkeit geschaffen, individuell und anonym eine Vorauswahl der Zielsysteme zu treffen. Inwieweit die Zahl der auswählbaren Kataloge dabei systemseitig beschränkt werden muß, wird eine gründliche Analyse zeigen müssen.

[23] Der Zugriff über Z39.50-Clients kann selbstverständlich auch von Benutzern erfolgen, die über ein solche Software verfügen. Eine Übersicht über die derzeit erhältlichen Z39.50-Clients bietet: Peter Evans: Z39.50 Standalone Client Software Review, in: BiblioTech Review; October 1997. Als Volltext unter: *http://www.gadgetserver.com/bibliotech/html/z3950clients.html*
In diesem Zusammenhang lohnt es sich, das Mozilla RDF /Z39.50 Integration Project zu beobachten, in dem ein frei verfügbarer WEB-Browser mit Z39.50-Client-Funktionalität entwickelt werden soll. Siehe. *http://www.mozilla.org/rdf/doc/z3950.html*

[24] Leider sind Profile, in denen die semantischen Suchmöglichkeiten der Z39.50-Server beschrieben sind, bislang eher die Ausnahme denn die Regel. Von den DBV-OSI-Partnern sind von den beteiligten Bibliotheken und Verbünden nur die Profile des Bayerischen Bibliotheksverbundes (BVB) und Der Deutschen Bibliothek im Internet publiziert siehe *http://www-opac.bib-bvb.de/subbvb/ordbvb/targetprofile.htm* und *http://www.ddb.de/partner/ddb_profile.htm*

[25] Die Grundlage der Spezifikation ist nachzulesen in Lohrum, Stefan; Schneider, Wolfram; Willenborg, Josef: Deduplication in KOBV, ZIB SC 99-05. - Juni 1999 Als Volltext unter: *http://www.kobv.de/docs/*

[26] Der Schwellwert 50 ist parametrisierbar.

[27] Welche Felder für die Dublettenprüfung herangezogen werden, obliegt der Parametrisierung. Für jedes Feld kann darüber hinaus festgelegt werden, welche Zeichen berücksichtigt werden (Normierung) ob trunkiert werden soll oder nicht (Trunkierung nach N-Zeichen oder nach einem definierten Delimiter).

[28] BIB1 Use Attribut- / Wertpaar

[29] Erste konzeptionelle Überlegungen zur Intergration von Fachinformationsystemen in den KOBV sind angestellt worden in dem Aufsatz von Lügger, Joachim: über Suchmaschinen, Verbünde und die Integration von Informationsangeboten. Internes Arbeitspapier. Oktober 1999.

[30] Penke, Kurt; Scharna, Dieter; Willenborg, Josef: Bibliotheksverzeichnisse zur Auswahl von Bibliotheksbeständen in verteilten Informationssystemen - Untersuchung und Einsatz im Kooperativen Bibliotheksverbund Berlin - Brandenburg (KOBV). In: Bibliotheksdienst 33,1999, Heft 10, S. 1667-1697.