

Bibliotheksdienst Heft 8/9, 96

Karlsruher Virtueller Katalog (KVK) Neue Dienstleistung im World Wide Web

Uwe Dierolf, Michael Mönnich

Bibliothekskataloge und -verbünde

Seit ca. zehn Jahren katalogisieren die großen Bibliotheken in Deutschland in die gemeinsamen Datenpools der Bibliotheksverbünde. Die derzeit größten sind:

- Bayerischer Verbund BVB (Bayern)
- Gemeinsamer Bibliotheksverbund GBV (Niedersachsen, Bremen, Hamburg, Schleswig-Holstein, Thüringen, Mecklenburg-Vorpommern, Sachsen-Anhalt)
- Südwestdeutscher Bibliotheksverbund SWB (Baden-Württemberg, Rheinland-Pfalz)
- Verbunddatenbank des Hochschulbibliotheksentrums HBZ (Nordrhein-Westfalen, Rheinland-Pfalz)
- Zeitschriftendatenbank des DBI (alle wissenschaftlichen Bibliotheken in Deutschland, nur Zeitschriften).

Die Verbünde enthalten umfangreiche Datenbanken mit Titelaufnahmen und Bestandsnachweisen von Büchern und Zeitschriften (Angaben aus den jeweiligen WWW-Servern, 29.7.96):

Verbund	Titel	Bestandsnachweise
BVB:	8.000.000	15.000.000
GBV:	6.400.000	11.000.000
HBZ:	6.500.000	12.000.000
SWB:	4.700.000	11.000.000
zusammen	25.600.000	49.000.000

Für eine umfassende Buchrecherche ist die Abfrage all dieser Kataloge notwendig. Da es sich bei allen um unterschiedliche EDV-Systeme handelt, ist die Nutzung auch über WWW nicht einfach. Es werden unterschiedliche Trunkierungszeichen, unterschiedliche Suchfelder und völlig unterschiedlich strukturierte Trefferlisten angeboten.

Neben den Bibliotheksverbänden bieten zahlreiche Bibliotheken selbst OPACs an, die den lokalen Bestand der Bibliothek enthalten. In Karlsruhe z. B.:

Bibliothek	Titel
Universitätsbibliothek:	320.000
Institutskatalog der Universität:	180.000
Badische Landesbibliothek:	400.000

Auch hier müssen bislang zur vollständigen Recherche der lokalen Bestände mehrere Einzelkataloge durchsucht werden.

WWW-Schnittstellen

Im Laufe des letzten Jahres haben zahlreiche Bibliotheken wie auch die Mehrzahl der Verbünde ihre Systeme mit WWW-Schnittstellen ausgerüstet, die es ermöglichen, Suchanfragen über HTML-Formulare (FORMS) einzugeben und mit HTML formatierte Trefferlisten zu liefern.

Die Benutzereingaben werden im HTML-Formular eindeutigen Suchaspekten zugeordnet und an ein Programm übergeben, das aus diesen Daten eine Datenbankanfrage formuliert und die Trefferlisten aufbereitet. Hierbei sind die CGI (Common Gateway Interface)-Konventionen einzuhalten.

Konzept des Karlsruher Virtuellen Kataloges (KVK)

Das Ziel des Karlsruher Virtuellen Katalogs (KVK) war es, über eine Suchanfrage in beliebigen Katalogen zugleich suchen zu können und ein Gesamtergebnis zu erhalten.

Hierzu wurde ein CGI-Programm erstellt, das

- die im KVK-Suchformular eingegebene Suchanfrage für mehrere Zielkataloge formuliert,
- die Anfrage dann parallel an alle Kataloge schickt,
- die einzelnen Trefferlisten sammelt und analysiert und
- zuletzt eine große Gesamttrefferliste in einem einheitlichem Format erstellt.

Die Gesamttrefferliste wird kontinuierlich aufgebaut. Sobald das Ergebnis eines Zielkatalogs vollständig vorliegt, wird es formatiert und sofort angezeigt. Das CGI-Programm des KVK-Systems kommuniziert direkt mit den WWW-Servern der Zielkataloge.

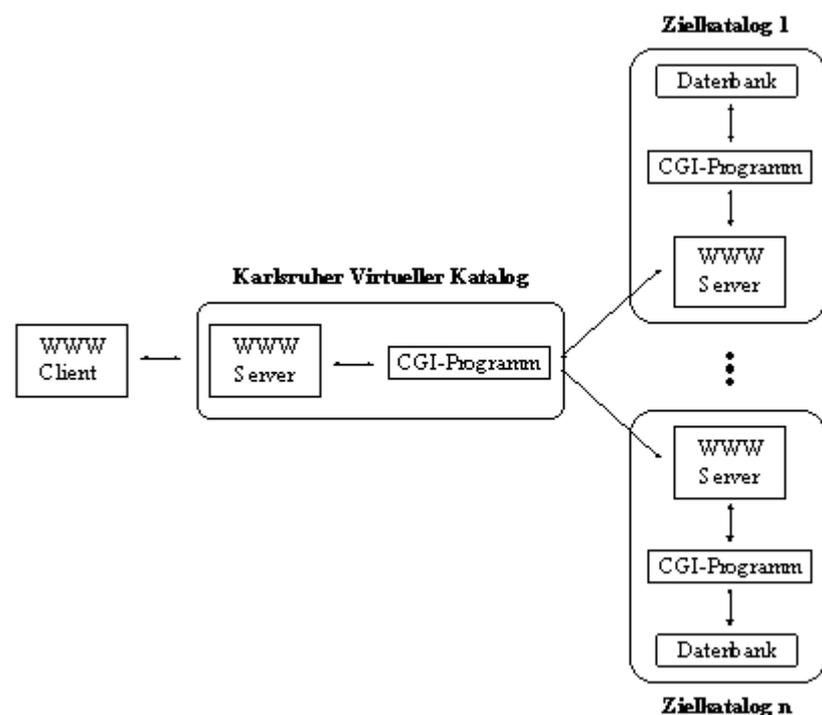


Abbildung 1: Architektur des Karlsruher Virtuellen Katalogs (KVK)

Ein ähnliches Verfahren wie das KVK-System nutzt auch die Internet-Suchmaschine MetaCrawler (<http://www.metacrawler.com>), um Suchmaschinen wie Lycos, Alta Vista und Infoseek parallel abzufragen.

Das KVK-System ist so ausgelegt, daß für jedes WWW-Suchinterface der einzelnen Zielkataloge eine Strukturbeschreibungsdatei geführt wird, die dem CGI-Programm als Eingabe dient. Sie beschreibt sowohl den Aufbau des Suchformulars als auch den Aufbau der Trefferlisten.

Die wichtigsten Strukturelemente eines HTML-Suchformulars (FORM) sind der WWW-Server und der

Name des CGI-Programms des Zielkatalogs, das die Suche durchführt. Außerdem müssen die Namen der Felder dem KVK-System bekannt sein. Die Feldnamen und -inhalte des KVK-Suchformulars werden auf die Felder des Zielkatalogs abgebildet. Hierdurch ist es möglich, für jeden Zielkatalog eine Suchanfrage in dessen Syntax zu erstellen.

```
<FORM METHOD=„POST“ ACTION=„http://www.ubka.uni-karlsruhe.de/hylib-bin/suche.cgi“>
Titel <INPUT name=„TI“ value=„,„>
Autor <INPUT name=„AU“ value=„,„>
<INPUT TYPE=„submit“ name=„Suchen“>
</FORM>
```

Abbildung 2: Beispiel für ein Suchformular

Die Kenntnis der Struktur der Trefferliste ist Grundlage für das Erstellen einer einheitlichen Gesamttrefferliste. Die wichtigsten Grundelemente, die hierbei erkannt werden müssen, sind der Kurztitel und der zugehörige Link (bzw. URL). Der Link verweist auf das CGI-Programm des Zielkatalogs und enthält i.d.R. die Suchanfrage für die Volltitelanzeige. Hierbei wird die Technik des Parsings auf Basis regulärer Ausdrücke eingesetzt (HTML-Parsing).

```
<UL>
<LI><A HREF=/hylib-bin/suche.cgi?opacdb=UB+Karlsruhe&nd=4141992>
Mastering MAKE / Clovis L. Tondo ; Andrew Nathanson ; Eden Yount , 1994</A>
<LI><A HREF=/hylib-bin/suche.cgi?opacdb=UB+Karlsruhe&nd=3630188>
Make und nmake : software management unter UNIX und MS-DOS. - 1. Aufl. -, 1994</A>
</UL>
```

R1: ^ Der Klammerausdruck entspricht dem Link
R2: ^(.*) Der Klammerausdruck entspricht dem Titel

Abbildung 3: Auszug aus einer Trefferliste und zugehörige reguläre Ausdrücke (R1,R2)

Derzeitige Realisierung

Über die Homepage der Universitätsbibliothek Karlsruhe

<http://www.ubka.uni-karlsruhe.de>

hat man Zugriff auf den Karlsruher Virtuellen Katalog, der Suchanfragen an folgende Kataloge ermöglicht:

- Südwestdeutscher Bibliotheksverbund (SWB) mit 4,7 Mio. Titeln aus Baden-Württemberg, Rheinland-Pfalz und Sachsen
- Bayerischer Verbund (BVB) mit 8 Mio Titeln
- Gemeinsamer Bibliotheksverbund (GBV) mit 6,4 Mio Titeln
- Verbunddatenbank des Hochschulbibliotheksentrums (HBZ) mit 6,5 Mio. Titeln aus Nordrhein-Westfalen und Rheinland-Pfalz) und
- die Karlsruher OLIX-OPACs mit zusammen 0,9 Mio. Titeln

Der Benutzer kann wählen, in welchen Katalogen gesucht werden soll; standardmäßig wird in allen recherchiert.

Universitätsbibliothek Karlsruhe

Karlsruher Virtueller Katalog (KVK)

[UB Karlsruhe](#) . . [Hilfe](#)

Welche Kataloge sollen bei der Suche berücksichtigt werden?

Verbünde und Buchhandel

- Südwestdeutscher Bibliotheksverbund
- Bibliotheksverbund Bayern
- Verbundkatalog Nordrhein-Westfalen
- Gemeinsamer Bibliotheksverbund
- Zeitschriftendatenbank (ZDB)
- KNO-K&V Buchkatalog
- Verzeichnis lieferbarer Bücher (VLB)

Bibliotheken

- Universitätsbibliothek Karlsruhe
- Instituts katalog Uni Karlsruhe
- Badische Landesbibliothek
- StOPAC (Stuttgarter Katalog)

Soll zusätzlich eine sortierte Liste ausgegeben werden? Ja Nein

Titel	<input type="text"/>	Schlagwort	<input type="text"/>
Autor	<input type="text"/>	ISBN	<input type="text"/>
Körperschaft	<input type="text"/>	ISSN	<input type="text"/>
Jahr	<input type="text"/>	Verlag	<input type="text"/>

Abbildung 4: Suchmaske des Karlsruher Virtuellen Kataloges

Man kann also in einem Datenbestand von ca. 20 Millionen Titeln recherchieren.

Der KVK liefert Trefferlisten mit formatierten Kurztitel, von denen durch einfaches Anklicken auf die Volltitelaufnahme und die Bestandnachweise des jeweiligen Zielkatalogs umgeschaltet werden kann. Zusätzlich erzeugt der KVK eine kumulierte Liste, welche die alphabetisch sortierten Kurztitel enthält.

Das System wurde an der UB Karlsruhe in Zusammenarbeit mit der Fakultät für Informatik im Rahmen einer Studienarbeit entwickelt. Es basiert auf der Skriptsprache TCL.

Virtueller Katalog als Alternative zu Normschnittstellen

Kann die bei KVK verwendete Technik eine Alternative zur Verwendung von Recherche-Schnittstellen auf Basis der normierten Protokolle SR bzw. Z39.50 (im folgenden als Normschnittstellen bezeichnet) darstellen?

Im Vergleich zu den Normschnittstellen bietet der bei KVK verfolgte Ansatz folgende Vor- und Nachteile:

Der KVK kann auch bei Systemen eingesetzt werden, die nicht über eine Normschnittstelle verfügen. Ein weiterer Vorteil liegt in der optimalen Ausnutzung der Hypertextfunktionalität des WWW durch das KVK-System. Jeder Treffer der vom KVK gelieferten Liste stellt einen Link auf einen WWW-Server des jeweiligen Zielkatalogs dar. Per Mausklick kann der Benutzer daher direkt weitere Dienste und

Funktionen (Bestandsnachweis, Ausleihe, Fernleihe, Abruf von multimedialer Zusatzinformation etc.) aufrufen.

Da Änderungen in den WWW-Schnittstellen der Zielkataloge Änderungen in den zugehörigen Strukturbeschreibungsdateien nach sich ziehen, kann nur eine begrenzte Anzahl von Katalogen mit vertretbarem Pflegeaufwand angeboten werden. Bei Verwendung von Normschnittstellen tritt dieses Problem nicht auf.

Weiterentwicklung

Die Erstellung der Strukturbeschreibungsdateien eines WWW-Suchinterfaces ist nicht sehr aufwendig. Es setzt lediglich Kenntnis über die Formulierung regulärer Ausdrücke zur Strukturerkennung voraus. Daher ist das System leicht erweiterbar, so daß weitere Kataloge bzw. Verbünde (z.B DBI-VK) in den KVK aufgenommen werden können, sobald diese ein WWW-Suchinterface anbieten.

Derzeit bestehen noch folgende Probleme:

- Bei manchen Zielkatalogen ist der Zugriff (wie z. B. beim Betatest-WWW-Interface der ZDB, <http://dbix01.dbi-berlin.de:8163/>) über Benutzername und Kennwort geschützt. Da diese Angaben interaktiv eingegeben werden müssen, ist die Integration in den Virtuellen Katalog nicht ohne Zusatzaufwand möglich. Derzeit kann in solchen Zielkatalogen nicht mittels des KVK recheriert werden. Es ist wünschenswert, wenn der Zugang zu einem WWW-OPAC (Online Public Access Catalog) auch in Zukunft frei zugänglich ist.
- Manche HTML-Suchformulare werden dynamisch generiert (s.o.: ZDB-WWW-Interface). Sie enthalten dabei meist einen variablen Anteil in Form einer unsichtbaren Sitzungsnummer HTML: `<INPUT TYPE=HIDDEN NAME="session_id" VALUE="0815">`.
- Diese Sitzungsnummer ist dann ein obligatorischer Parameter bei allen Anfragen an den Zielkatalog. Das Konzept des KVK müßte dahingehend erweitert werden, daß vor dem Versand einer Suchanfrage die Sitzungsnummer ermittelt wird. Derzeit ist dies jedoch noch nicht realisiert.
- Änderungen an den HTML-Suchformularen werden derzeit nicht automatisch erkannt. Es sind daher Mechanismen zu realisieren, die Änderungen automatisch erkennen, z. B. ein regelmäßiger Abgleich der aktuellen Suchformulare mit den Strukturbeschreibungsdateien.