

## **E-LIB Bremen – Automatische Empfehlungsdienste für Fachdatenbanken im Bibliothekskatalog / Metadatenpools als Wissensbasis für bestandsunabhängige Services**

**Martin Blenke, Rachel Ellis, Elmar Haake**

Im Vergleich zur rasanten Entwicklung von Suchmaschinenteknik außerhalb des Bibliotheksumfelds haben sich die weitgehend Printmedien-basierten OPACs der Bibliotheken seit dem Jahr 2000 nur wenig inhaltlich weiterentwickelt. Aus diesem Grunde sind derzeit an vielen Stellen Bibliotheksinitiativen zu beobachten, die im Rahmen von Eigenentwicklungen neue Wege der Katalogsuche und Trefferpräsentation ausprobieren<sup>1</sup>. Dabei spielen Anregungen aus der nichtbibliothekarischen Suchmaschinenszene und dem Kontext des Web 2.0 eine wichtige Rolle.

Im Gegensatz zum einfachen Identifizieren und Anzeigen von Trefferlisten im klassischen OPAC erlauben es neue Indexierungs- und Präsentationstechniken, die wertvollen und aufwendig erzeugten bibliografischen Daten des Suchkontextes nutzerfreundlicher darzustellen. Außerdem können im Verlauf des Such- und Rankingvorgangs Analysen der Metadaten der gefundenen Titel erfolgen, auf deren Basis die Entwicklung neuer Katalogfunktionen möglich ist. Ein Beispiel für einen neuen Dienst, der nach diesem Prinzip aufgebaut werden kann, ist die automatische Empfehlung von bibliografischen Fachdatenbanken im Rahmen der Katalogsuche in der E-LIB Bremen.

### **E-LIB Bremen – Suchmaschinenteknik für den Bibliothekskatalog**

Die Elektronische Bibliothek (E-LIB) der SuUB Bremen<sup>2</sup> bietet bereits seit 2002 einen einheitlichen Katalog elektronischer und gedruckter Medien auf Basis von Suchmaschinenteknik. Im Sinne eines „One-Stop-Shop-Portals“ werden alle verfügbaren Medien und Dienste der Bibliothek unter einer Suchoberfläche an-

- 1 Lossau, Norbert: Search engine technology and digital libraries: libraries need to discover the academic internet – D-lib magazine, 10, 6 (2004), <http://dx.doi.org/10.1045/june2004-lossau>  
Eine Auswahl weiterer neuer Katalogprojekte in Deutschland – einige davon noch Prototypen, andere bereits seit längerer Zeit produktiv im Einsatz:  
XOPAC – UB Karlsruhe, <http://www.ubka.uni-karlsruhe.de/katalog/>  
Heidi – UB Heidelberg, <http://katalog.ub.uni-heidelberg.de/>  
OpenBIB – UB Köln, <http://www.openbib.org/>  
Beluga – SUB Hamburg, <http://beluga.sub.uni-hamburg.de/>  
Suchkiste – GBV Verbundzentrale, <https://finden.nationallizenz.de/>
- 2 E-LIB Bremen Suchmaschine, <http://suche3.suub.uni-bremen.de/>

geboten, um die übliche heterogene Nachweissituation der verschiedenen Medientypen soweit wie möglich aufzulösen. Das System verfolgt damit das Ziel, alle Bestände, die für Bremer Nutzer lokal vor Ort und extern verfügbar sind, in einem einzigen Retrievalsystem zur Verfügung zu stellen. In diesem von der Bibliothek kontrollierten Suchraum der Bremer E-LIB finden sich derzeit ca. 22 Mio. Nachweise – davon sind 80% elektronische Volltexte. Die Nachweisbasis wird durch Integration weiterer Metadatenansammlungen kontinuierlich erweitert. Die E-LIB Bremen versteht sich gleichzeitig als Produktivsystem<sup>3</sup> und Entwicklungsumgebung für neue Services zur Unterstützung von Retrievalvorgängen. Die eingesetzte Suchmaschinenteknik basiert auf einem lokal gepflegten Gesamtindex von heterogenen, aggregierten Metadaten. Die Organisation dieser Daten erledigt das System CiXbase<sup>4</sup>, ein Rechercheassistent unter Nutzung computerlinguistischer Verfahren. CiXbase ist eine Weiterentwicklung aus dem OSIRIS-Projekt<sup>5</sup> und wird nicht nur in Bremen, sondern u.a. auch in der Schweiz an der ETH Zürich (NEBIS) eingesetzt<sup>6</sup>.

Typische Vorteile einer Suchmaschine unter Nutzung eines lokalen Gesamtindex sind die hohe Antwortgeschwindigkeit<sup>7</sup>, die homogene Trefferpräsentation und eine ausgewogene Gewichtung aller Ressourcen. Die lokale Datensammlung eröffnet darüber hinaus Möglichkeiten, die vorhandenen Nachweise als Wissensbasis für neue Verfahren der Treffermengenanalyse zu nutzen, die während der Suche auf der Menge der Resultate prozessiert werden können. Dazu gehören z.B. klassische Drill-Down-Techniken, die nachträglich eine formale und inhaltliche Filterung erzielter Treffersets erlauben und die bereits in vielen modernen Suchmaschinen zum Standard gehören. Eine wichtige neue Strategie der E-LIB ist die kontextsensitive Unterstützung der Nutzerinnen und Nutzer durch verschiedene

- 3 Derzeit werden ca. 8.000 bis 13.000 Suchen täglich registriert, die von ca. 40.000 unterschiedlichen Rechnern im Monat durchgeführt werden (Stand April 2009).
- 4 CiXbase: <http://cixbase.dyndns.org/CiXbase/cixdocs/>
- 5 Zillmann, Hartmut: *Liber Quarterly* 10, S. 335–341 (2000), <http://webdoc.gwdg.de/edoc/aw/liber/lq-3-00/lq-3-00-s335b341.pdf>
- 6 NEBIS ETH Zürich: [http://www.nebis.ch/index\\_e.html](http://www.nebis.ch/index_e.html)  
CAMbase: Thomas Ostermann u.a. CAMbase – A XML-based bibliographical database on Complementary and Alternative Medicine (CAM) – Biomedical Digital Libraries, 4:2 / doi:10.1186/1742-5581-4-2 (2007) <http://www.bio-diglib.com/content/pdf/1742-5581-4-2.pdf>
- 7 Die Antwortzeiten vergleichbarer Systeme, die auf Metasuchtechniken nach dem Prinzip des "federated searching", d.h. ohne Nutzung einer lokalen Indexbasis operieren, sind oft kritisch und in der Wahrnehmung der Kunden zu lang. Siehe auch: Rochkind, J. (Meta)search Like Google. The time has come for libraries, too, to negotiate for rights to index full text (2007) – *Library Journal*: <http://www.libraryjournal.com/article/CA6413442.html>

Zusatzfunktionen, die auf Basis einer statistischen und computerlinguistischen Auswertung der Metadaten von Treffer sets angeboten werden.

### **Kontextsensitive Integration fachspezifischer Angebote im Katalog**

Geprägt durch die praktische Erfahrung mit Internet-Suchmaschinen<sup>8</sup> tritt die Katalogsuche als zentraler und wichtigster Bibliotheksservice immer mehr in den Fokus der Nutzer, die über die Mediensuche ein möglichst integriertes Angebot aller Dienste erwarten. Die Wahrnehmung anderer Angebote, die neben dem OPAC über die Bibliothekswebseiten angeboten werden, ist dagegen in der Regel deutlich reduziert.

Während die Recherche mit dem lokalen Bibliothekskatalog daher relativ einfach zu vermitteln ist, stellt die Vielfalt der unterschiedlichen externen Informationsmittel und deren spezifische Zugangskonditionen ein großes Hindernis bei der effizienten Nutzung dieser Quellen dar.

Die Prozesskette der Literaturrecherche, die im lokalen Katalog beginnt und durch überregionale Bibliotheksdienste, bibliografische Fachdatenbanken und zahlreiche weitere Instanzen verlaufen kann, wird durch die vorhandenen Onlineangebote der Bibliotheken meist nur unzureichend unterstützt.

Insbesondere fachspezifische Dienste sind in den meisten Bibliothekskatalogen unterentwickelt. Sie werden bisher in der Webpräsenz der Bibliotheken üblicherweise auf eigenständigen Informationsseiten außerhalb der Retrievalumgebung präsentiert. Von den Nutzerinnen und Nutzern wird somit erwartet, das Vorhandensein dieser Angebote abseits der zentralen Suche zu antizipieren und aktiv aufzusuchen.

Ein Beispiel für fachspezifische Dienste, die vom lokalen OPAC isoliert angeboten werden, ist in der Regel das Gesamtangebot der bibliografischen Fachdatenbanken einer Bibliothek. Gleichzeitig wird die unter Kunden geringe Kenntnis und Nutzungsfrequenz der teuren lizenzpflichtigen Fachdatenbanken mit Recht bereits seit längerer Zeit problematisiert<sup>9</sup>.

---

8 Keane, Mark T. u.a.: Are people biased in their use of search engines? – Communications of the ACM, 51, 2, S. 49–52 (2008), <http://doi.acm.org/10.1145/1314215.1314224>

9 Schüller-Zwierlein, Andre: Die Vermittlung der Schlüsselqualifikation Informationskompetenz an der LMU München : ein Lagebericht. 2006 – München : LMU. – <http://epub.ub.uni-muenchen.de/1349/>; Tietze-Netolitzky, Almut u.a.: Bankenpleiten? Datenbanken und Benutzernachfrage : eine Umfrage in den geistes- und sozialwissenschaftlichen Teilbibliotheken der UB München – *Bibliotheksdienst*, 37, 3, S. 302–319 (2003); Klatt, Rüdiger (Hrsg.): Elektronische Information in der Hochschulausbildung : innovative Mediennutzung im Lernalltag der Hochschulen, Opladen 2001 sowie SteFi Studie aus dem Jahr 2001, <http://www.stefi.de/>

Liefert eine thematische Recherche im lokalen Bibliothekskatalog unbefriedigende Resultate, ist in vielen Fällen die Nutzung einer geeigneten fachlich spezialisierten bibliografischen Datenbank zu empfehlen. Der Kunde erhält für den damit notwendigen Medienbruch beim Übergang vom lokalen Katalog zum externen Datenbankanbieter jedoch kaum Unterstützung durch konventionelle OPAC-Systeme. Versuche zur Implementierung von Applikationen, die auf Basis von hinterlegtem Expertenwissen Fachdatenbankempfehlungen zur Unterstützung der Literatursuche generieren können, sind in der Literatur beschrieben, haben allerdings bisher kaum Verbreitung gefunden<sup>10</sup>.

Es ist daher aus unserer Sicht naheliegend, die Informationen über weitere verfügbare fachliche Ressourcen direkt an der Stelle und in der Situation anzubieten, in der ein Nutzer ein fachliches Interesse im Rahmen einer Katalogsuche äußert.

### **Expertensystem zur Empfehlung von bibliografischen Fachdatenbanken**

Eine Strategie der E-LIB-Suchmaschine besteht darin, das thematische Umfeld der Anfrage anhand der vom Nutzer gewählten Suchbegriffe zu analysieren und daraufhin geeignete, auf das Fach zugeschnittene Dienste direkt im Katalog anzubieten. Dazu wird jedes Suchergebnis zunächst computerlinguistisch und statistisch ausgewertet und daraus anschließend unter anderem eine grobe Fachzuordnung anhand der Bremer Systematik vorgenommen. Jeder Anfrage können eine oder mehrere fachliche Schwerpunkte zugeordnet werden, falls das Suchergebnis eine statistisch relevante Anzahl systematisch erschlossener Treffer aufweist. Ein auf diesem Prinzip basierender Fachprofil-Dienst kann in vielfältiger Weise zum Aufbau erweiterter Katalogfunktionen verwendet werden. Prinzipiell bietet sich die Anzeige aller fachspezifischen Informationen an, die für die Nutzer in der jeweiligen Recherchesituation als sinnvoll erachtet werden und den weiteren Prozess der Mediensuche unterstützen können, z.B.:

- Empfehlenswerte bibliografische Fachdatenbanken
- Informationen und Kontakt zu fachlichen Ansprechpartnern
- Aktuelle, fachbezogene Neuigkeiten
- Schulungsangebote der jeweiligen Fächer
- Einstieg zur Online-Fachsystematik einer Disziplin.

In der E-LIB Bremen wird der Fachprofil-Service bisher zur Anzeige von geeigneten bibliografischen Fachdatenbanken und zum Einstieg in die Bremer Online-Fachsystematiken genutzt.

<sup>10</sup> Ein halbautomatisches Empfehlungssystem zur geführten Auswahl von Fachdatenbanken ist z.B. das System BibTutor. Das vom BMBF geförderte Projekt BibTutor und ähnliche Angebote mit didaktischem Anspruch zur Vermittlung von Informationskompetenz sind recht komplex zu bedienen und operieren ebenfalls als isolierte Anwendung. <http://www.dfki.uni-kl.de/bibtutor/>

Die Metadaten der ca. 200 in Bremen verfügbaren Fachdatenbanken werden in einer eigenen Tabelle vorgehalten. Dabei erfolgt zusätzlich eine intellektuelle Gewichtung der Titel durch die Fachreferate. Wichtige lizenzierte Angebote werden gegenüber kleineren Spezialdatenbanken bevorzugt behandelt.

Bei einer Recherche in der Suchmaschine wird zunächst das Fachprofil der Recherchebegriffe durch Analyse der Metadaten der Treffermenge ermittelt. Das System präsentiert anschließend die Empfehlungen zu geeigneten bibliografischen Fachdatenbanken in einer Zusatzbox rechts neben der Trefferliste (siehe Abb. 1: Suchbeispiel „Englischunterricht“).

The screenshot shows a search interface with the following elements:

- Search Bar:** Contains the text 'englischunterricht' and a 'Suche starten' button with a timer '(0,01 sec)'.
- Filters:** Radio buttons for 'nur Zeitschriftentitel', 'nur elektronische Zeitschriftentitel', 'nur Bücher', and 'nur E-Books'.
- Results Summary:** '870 Titel: englischunterricht -> [Bücher]', 'Modus: Themensuche (Sortierung: Relevanz)', and a link to 'Einfache Suche (Sortierung: Jahr)'.
- Search Results (1-5):**
  - 1.** Englischunterricht : Grundlagen und Methoden einer handlungsorientierten Unterrichtspraxis UTB für Wissenschaft ; 1540, Anglistik / Bach, Gerhard - 3., vollst. überarb. und verb. Aufl. , 2003. Exemplar(e): a pae 612 m/235(3); a pae 612 m/235(3)a; a pae 612 m/235(3)b; DL Wendt td 6293; h pae 612 m/235(3)c; a pae 612 m/235(3)d; a pae 612 m/235(3)e; a pae 612 m/235(3)f; a pae 612 m/235(3)g; a pae 612 m/235(3)h; a pae 612 m/235(3)i; a pae 612 m/235(3)j; a pae 612 m/235(3)k; a pae 612 m/235(3)l; a pae 612 m/235(3)m; DL Abendroth te 1059. Export in [RefWorks] [Bibsonomy] [Biblex] [EndNote] [RefMan(RIS)]
  - 2.** Englischunterricht : Grundlagen und Methoden einer handlungsorientierten Unterrichtspraxis UTB für Wissenschaft, Uni-Taschenbücher ; 1540 / Bach, Gerhard - 2., überarb. und erw. Aufl. , 1996. Exemplar(e): h pae 612 m/235(2); a pae 612 m/235(2)a; a pae 612 m/235(2)b; pae 612 m/235(2)c; a pae 612 m/235(2)d. Export in [RefWorks] [Bibsonomy] [Biblex] [EndNote] [RefMan(RIS)]
  - 3.** Englischunterricht : Grundlagen und Methoden einer handlungsorientierten Unterrichtspraxis Uni-Taschenbücher ; 1540, Anglistik; UTB ; 1540, Anglistik / Bach, Gerhard ; Timm, Johannes-Peter , 1989. Exemplar(e): a pae 612 m/235. Export in [RefWorks] [Bibsonomy] [Biblex] [EndNote] [RefMan(RIS)]
  - 4.** Englischunterricht für alle : Beiträge zur Didaktik und Methodik des Englischunterrichts in Lerngruppen mit schwierigen Voraussetzungen / Hellwig, Karlheinz; Sauer, Helmut , 1984. Exemplar(e): a pae 612 m/141. Export in [RefWorks] [Bibsonomy] [Biblex] [EndNote] [RefMan(RIS)]
  - 5.** Englischunterricht : Literatur 5-10 U- und S-Pädagogik, Praxis und Theorie des Unterrichtens / Hunfeld, Hans , 1982. Exemplar(e): a pae 612 6/602. Export in [RefWorks] [Bibsonomy] [Biblex] [EndNote] [RefMan(RIS)]
- Teilaspekte des Themas: englischunterricht:**
  - Anglistik Aufsatzsammlung Deutschunterricht
  - Drama Englisch Englischlehrer English language
  - English language Study and teaching Foreign speakers
  - Erwachsenenbildung Französischunterricht
  - Fremdsprache Fremdsprachenlernen
  - Fremdsprachenunterricht Grammatik Großbritanien
  - Grundschule Gymnasium Hauptschule
  - Landeskunde Lehrmittel Lernspiel Literatur
  - Literaturunterricht Schulbuch Sekundarstufe
  - Sekundarstufe 1 Sekundarstufe 2 Spiel USA
  - Wortschatz
- Datenbankempfehlung:**
  - Pädagogik:**
    - 👉 Datenbank Bildung: FIS Bildung
    - 👉 Datenbank Bildung: ERIC Database
  - Anglistik:**
    - 👉 MLA - Datenbank der Modern Language Association
    - 👉 Annual Bibliography of English Language Literature (ABELL)
    - 👉 Online Contents Anglistik

Abb. 1: Ausgabe der Fachdatenbankempfehlung als Zusatzbox neben der Trefferliste für die Beispielsuche „Englischunterricht“. Die Empfehlung erfolgt aufgrund einer intellektuellen Zuordnung der wichtigsten Systeme für das jeweilige Fach<sup>11</sup> (Die Drill-Down-Funktionen der Suchmaschine fehlen in der Abbildung)

11 Die Zuordnung erfolgt derzeit auf der obersten Hauptklasse der Bremer Systematik (also z.B. „Chemie“). Sie könnte bei Bedarf aber beliebig feiner angewendet werden z.B. unter Differenzierung in größere Fachcluster wie „Anorganische Chemie“ und „Organische Chemie“ etc.

In der Box „Datenbankempfehlung“ werden bei den erkannten Fächern die zugeordneten Datenbanken mit der höchsten Priorität angezeigt und die Möglichkeit geboten, die Suche dort direkt fortzusetzen. Ein Zugriff auf externe Verzeichnisse von Fachdatenbanken ist in der Mehrzahl der Fälle nicht mehr notwendig.

Dies setzt jedoch voraus, dass die neben der Trefferliste angeordneten Zusatzdienste auch entsprechende Beachtung finden. Die Positionierung dieser Angebote auf der Seite der Ergebnispräsentation hat einen direkten Einfluss auf deren Wahrnehmung und ist derzeit Gegenstand intensiver Untersuchungen<sup>12</sup>. Das Screendesign der E-LIB Suche befindet sich daher in einem kontinuierlichen Prozess der Überarbeitung.

Zukünftig soll der Dienst mehrstufig erweitert werden. In der nächsten Phase der Weiterentwicklung sollen die Datenbankverweise so gestaltet werden, dass die ursprünglich formulierte Suche direkt an die entsprechenden externen Systeme übergeben werden kann, falls die jeweiligen Zielsysteme diese Möglichkeit technisch bieten. Weitere Ausbauperspektiven ergeben sich durch die Integration von Metasuchtechniken: falls die bibliografischen Datenbanken entsprechende Schnittstellen bereitstellen, ist zusätzlich auch die Angabe der zu erzielenden Trefferzahl für die ursprünglich formulierte Anfrage in den empfohlenen Systemen möglich<sup>13</sup>.

Ein solcher Dienst sollte prinzipiell auch in anderen Katalogsystemen implementierbar sein, in denen eine ausreichende Menge von systematisch erschlossenen Titeln verfügbar ist.

### **Bibliografische Daten als Wissensbasis?**

Das Prinzip der Analyse von Treffermengen ist nicht nur auf die hier beschriebene fachliche Zuordnung von Anfragen beschränkt. Weitere Dienste, die das System ebenfalls in der rechten Spalte der Trefferpräsentation anzeigt, erlauben etwa eine Tag Cloud-Visualisierung des begrifflichen Umfeldes (siehe Abbildungen) oder des zeitlichen Publikationsverlaufs des Themas der Anfrage. Bei Autoren-

12 Cutrell, Edward u.a.: What are you looking for? : an eye-tracking study of information usage in web search – CHI,07: Proceedings of the SIGCHI conference on Human factors in computing systems, S. 407–416. (2007); <http://doi.acm.org/10.1145/1240624.1240690>; Jung, Seikyung u.a.: LibraryFind: System design and usability testing of academic metasearch system – Journal of the American Society for Information Science and Technology 59, 3, S. 375–389 (2008); <http://dx.doi.org/10.1002/asi.20749>

13 Eine automatische Priorisierung von geeigneten Fachdatenbanken anhand der jeweils für eine Anfrage zu erzielenden Trefferzahlen wird derzeit ebenfalls geprüft. Jedoch ist dieser rein quantitative Parameter häufig nicht ausreichend für eine adäquate Zuordnung, da die Größe der Datenbasis der externen Zielsysteme stark schwanken kann.

recherchen ist die Präsentation von Koautorenclustern und eine Auflistung von Arbeitsgebieten der Personen möglich.

Diese Zusatzdienste, die neben der eigentlichen Suchfunktion in der E-LIB Bremen zur Verfügung stehen, sind vollständig modular aufgebaut und können daher als Teil-Module im Sinne eines Webservice<sup>14</sup> unabhängig in Systeme anderer Anbieter integriert und dort weitergenutzt werden.

Um die Nachnutzung solcher Analysefunktionen in Drittsystemen grundsätzlich zu ermöglichen, ist es jedoch notwendig, in einem vorgegebenen Rahmen allgemeingültige Resultate zu gewährleisten. Wenn die vorhandene lokale Datenbasis eine entsprechende Größe erreicht, fachlich umfassend und gleichmäßig verteilt ist, kann man davon ausgehen, dass die Dienste einen gewissen Grad der Unabhängigkeit vom lokal vorhandenen Bestand erreichen können und damit universell in anderen Systemen nachnutzbar sind.

The screenshot shows a search interface with a search bar containing 'malaria' and a 'Suche starten' button. Below the search bar, there are filters for 'nur Zeitschriftentitel' and 'nur elektronische Zeitschriftentitel'. The search results are displayed in a list format, with the first five results visible. On the right side, there are two sidebars: 'Teilaspekte des Themas:' which lists related terms like 'Anopheles', 'Chloroquine', and 'Plasmodium falciparum'; and 'Datenbankempfehlung:' which lists 'Medline (via PubMed)', 'Cinahl', and 'PreCinahl'.

Abb. 2: Ausgabe der Fachdatenbankempfehlung und des Begriffsumfelds für das Suchbeispiel „Malaria“

14 Ein Webservice ist eine Anwendung, die von anderen Applikationen über eine eindeutige Webadresse aufgerufen werden kann und Daten XML-kodiert zurückliefert.

Im Testbetrieb wird der Bremer Fachprofil-Service derzeit vom Hamburger Katalogprojekt Beluga genutzt, um dort ebenfalls anfragebezogene Datenbankempfehlungen zu generieren<sup>15</sup>. Dabei wird die Suchanfrage an das Beluga-System im Hintergrund von der E-LIB analysiert und als Antwort die ermittelte fachliche Zuordnung zurückgegeben. Auf Basis dieser Information kann Beluga anschließend eine auf den Hamburger Standort zugeschnittene Datenbankempfehlung präsentieren. Erste Tests zeigen, dass die Fachzuordnungen, die auf Basis des Bremer Metadatenmaterials generiert werden können, auch in einem vergleichbaren anderen universitären Umfeld zu verwertbaren Zuordnungen führen. Die modulare Bereitstellung und gegenseitige Nutzung von Diensten nach dem Prinzip von Webservices kann damit neue Formen der Bibliothekskooperation zwischen einzelnen Standorten eröffnen. Es bleibt zu hoffen, dass mehr Einrichtungen eigene Entwicklungskompetenzen nutzen, um an solchen Service-Netzwerken teilzunehmen.

---

15 Christensen, Anne: beluga 0.7 online! in: Weblog des Projektes Beluga der SUB Hamburg;  
<http://beluga-blog.sub.uni-hamburg.de/blog/2009/04/02/beluga-07-online/> (2009)