Suche im Internet und neue Dienste der Universitätsbibliothek

Dr. Michael Voß (Universitätsbibliothek), mvoss@ub.hu-berlin.de

Kurzfassung: Die Internetsuche mit Suchmaschinen lässt Datenbanken und kostenpflichtige Inhalte aus. Die unterschiedlichen Oberflächen bei verschiedenen Datenbankbetreibern und die fehlende Möglichkeit, gleichzeitig bei verschiedenen Anbietern suchen zu können, sollen mit der Software MetaLib [1] behoben werden. Werden bei einer Suche in Datenbanken (direkt oder mit Hilfe von MetaLib) Ergebnisse erzielt, kann der Zugriff auf die Dokumente selbst mit Hilfe von SFX[1] unabhängig von deren physikalischer Form vermittelt werden (Online-Zugriff, Bestellung zur Ausleihe, ...). SFX erlaubt auf der Basis von Open-URL die Verknüpfung unterschiedlicher Datenbasen entsprechend den Vertragsbedingungen an der Universität.

Einleitung

Seit jeher haben die Bibliotheken den Auftrag, Medien zu sammeln, zu erfassen und der Benutzung zugänglich zu machen. In der Antike waren es Tontafeln und Papyrusrollen, später gedruckte Bücher und Zeitschriften und seit einigen Jahren zusätzlich elektronische Dokumente und Datensammlungen, die im Internet zugänglich sind oder von der Universitätsbibliothek käuflich erworben werden.

Seit 13 Jahren wird an der Universitätsbibliothek der Bestand an Büchern und Zeitschriften in einem Online-Katalog (OPAC) schrittweise erfasst. Zurzeit sind ca. zwei Mio. Titel elektronisch nachgewiesen und damit über das Internet recherchierbar. Damit haben wir ein Drittel unseres Gesamtbestandes von ca. sechs Mio. Medieneinheiten eingearbeitet.

Mit Beginn der 90er Jahre hat die Universitätsbibliothek begonnen, die damals gerade verstärkt aufkommenden elektronischen Medien innerhalb der Universität zur Benutzung zur Verfügung zu stellen. Waren es damals nur vereinzelte Datenbankrecherchen, die an einzelnen Plätzen und unter Anleitung durchgeführt werden konnten, oder die innerhalb des Universitäts-Rechnernetzes verfügbaren CD-ROM-Datenbanken, so sind heute über 8 000 elektronische Zeitschriften und eine Vielzahl von freien und kostenpflichtigen Datenbanken über die Nachweise der Universitätsbibliothek erreichbar.

Die Universitätsbibliothek hat in ihre Kataloge bisher die elektronischen Ressourcen nur zum Teil aufgenommen. So werden zwar Datenbanken oder elektronische Zeitschriften wie die Papierausgaben katalogisiert, wenn sie käuflich erworben werden, aber viele Internet-Ressourcen blieben bisher unbeachtet. Sieht man sich aber die Referenzen von Veröffentlichungen und Studienarbeiten an, so werden Internet-Ressourcen immer häufiger zitiert. Die Frage der Beschaffung, Katalogisierung und Bereitstellung dieser Informationen und Ressourcen wird für die Universitätsbibliothek immer wichtiger, damit sie bei der Veränderung der Arbeitsweise der Wissenschaftler und Studenten weiterhin ihrer Aufgabe gerecht werden kann.

Wenn es um die Suche von Informationen im Internet geht, kommt man um die verschiedenen Internet-Suchmaschinen (Google, AltaVista, Lycos, AllTheWeb, ...) nicht herum. Diese Art von Suchmöglichkeiten ist aber im Wesentlichen auf statische und frei zugängliche WWW-Seiten beschränkt. Sucht man mit Google zum Beispiel nach »Goethe und Farbenlehre«, findet man eine Vielzahl von WWW-Seiten – auch an der Humboldt-Universität –, aber keinen Hinweis, dass die Universitätsbibliothek entsprechende Exemplare des goetheschen Werkes besitzt. Gar nicht berücksichtigen kann eine solche allgemeine Suchmaschine die vertraglichen Beziehungen, die die Universität oder die Universitätsbibliothek mit Datenbankanbietern eingeht. Viele der im Internet gespeicherten Informationen sind nur mit entsprechenden Lizenzverträgen nutzbar. Die Universitätsbibliothek kauft pro Jahr in erheblichem Umfang Lizenzen für den Zugriff auf Online-Datenbanken und zu Volltexten von

1 SFX ist eine Abkürzung aus der Filmbranche und bezeichnet »»special effects«, vom Erfinder des Standards Open-URL genutzt, um zu kennzeichnen, dass mit einer Open-URL spezielle Dienste vermittelt werden können.

elektronischen Zeitschriften und anderen elektronischen Veröffentlichungen.

Datenbanken und Zugriff auf die Dokumente

In Datenbanken werden in großem Umfang Informationen über wissenschaftlich relevante Veröffentlichungen gespeichert und laufend aktualisiert. Den Zugriff auf einige dieser Datenbanken hat die Universitätsbibliothek für die Mitglieder der Humboldt-Universität lizenziert. Unter anderem sind hier Web of Science oder Current Contents Connect zu nennen. Wenn in diesen oder anderen Datenbanken recherchiert wird, kann mit den Treffern (Suchergebnissen) nur zum Teil unmittelbar weitergearbeitet werden. In einigen Fällen funktionieren die mit den Treffern gelieferten Links (z. B. zum Volltext), häufig landet man dann allerdings auf Seiten, auf denen ein kostenpflichtiges Login verlangt wird. Dies liegt zum Teil daran, dass die Universitätsbibliothek zu diesem Volltext wirklich keinen Lizenzvertrag abgeschlossen hat. Zum Teil liegt es aber daran, dass die Universitätsbibliothek bei einem anderen Anbieter Zugangslizenzen erworben hat. Für den Leser der Universitätsbibliothek, der ja heutzutage solche Recherchen an seinem Arbeitsplatz durchführt, ist dies aber nicht durchsichtig oder nur schwer nachvollziehbar. Sein Bestreben ist es, das Dokument, zu dem er die Metadaten in einer Datenbank gefunden hat, möglichst schnell zur Verfügung zu haben.

Zudem gibt es zu einem Fachgebiet unterschiedliche Datenbanken, die nacheinander durchsucht werden müssen, wenn man zu einer Fragestellung in einer Datenbank keine passenden Informationen gefunden hat. Dabei muss ein Wissenschaftler oder Student natürlich zunächst einmal wissen, welche der unzähligen Internet-Ressourcen wirklich relevante und originale Quellen nachweisen oder besitzen. Da viele dieser Internet-Ressourcen neu entstehen, ihren Schwerpunkt ändern oder zum Teil auch veralten, ist diese Aufgabe hinreichend schwierig. Das nächste Problem sind die unterschiedlichen Oberflächen der Datenbanken, in die man sich einarbeiten muss, um effektiv zu suchen und die richtigen Ergebnisse bei der Suche zu erzielen.

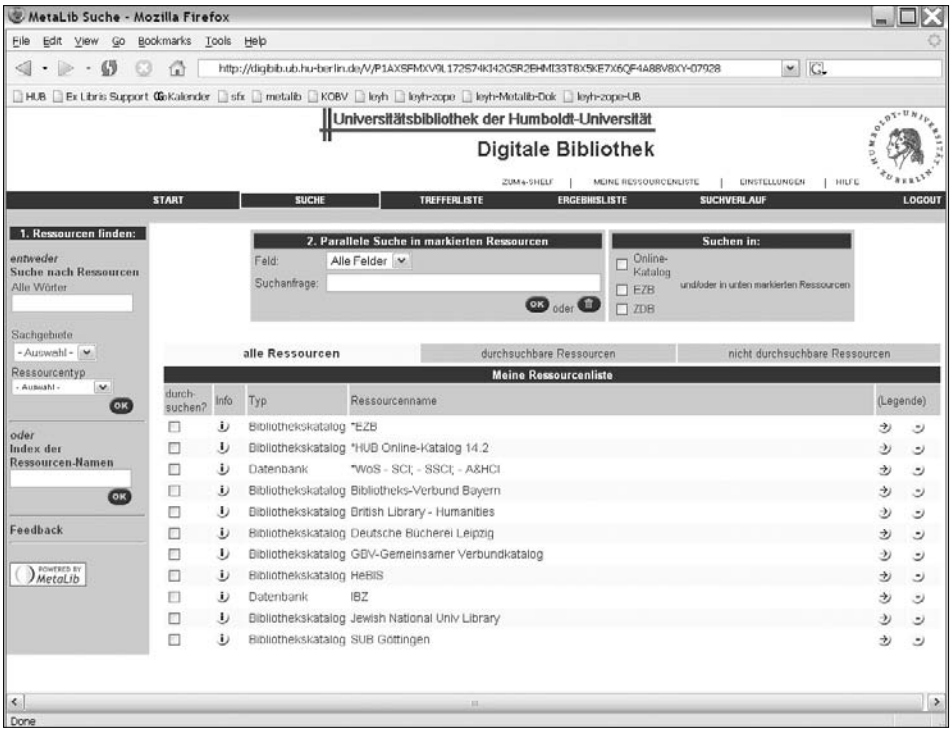

Abb. 1: Index von katalogisierten Internet-Ressourcen in MetaLib

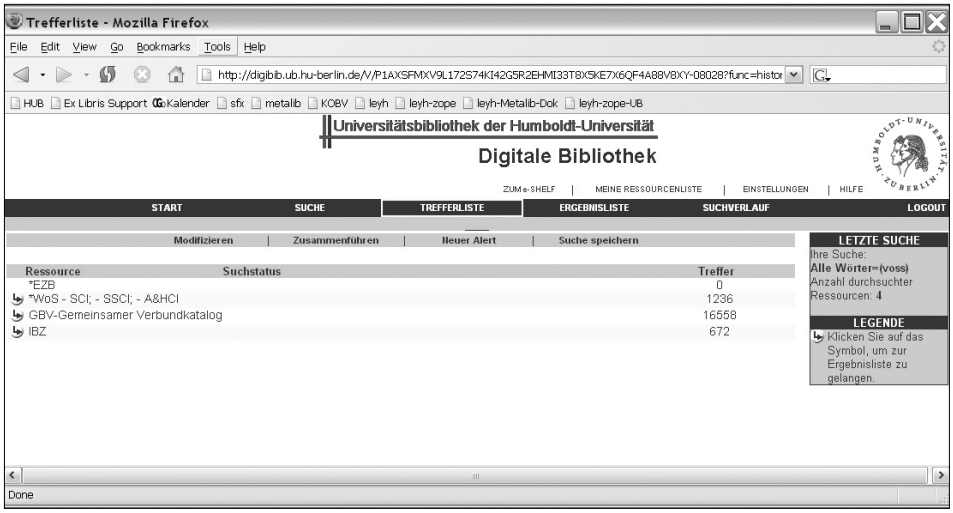

Abb. 2: Trefferanzeige von MetaLib

Die Digitale Bibliothek – MetaLib

Die Universitätsbibliothek hat sich diesen Entwicklungen bereits seit einiger Zeit gestellt und den Lesern der Universitätsbibliothek und allen Mitgliedern der Humboldt-Universität Unterstützung gewährt.

In der Vergangenheit erfolgte die Unterstützung im Wesentlichen durch Schulungen, Link-Listen und natürlich durch die Beschaffung kostenpflichtiger Inhalte durch Lizenzverträge. Dies reicht heute nicht mehr aus.

Mit der Beschaffung des Produktes MetaLib wird die Universitätsbibliothek beginnen, neben dem Online-Katalog, in dem weiterhin alle von der Universitätsbibliothek erworbenen Dokumente katalogisiert werden, einen Spezialkatalog von Internet-Ressourcen zu erstellen. In diesem Katalog werden die von der Universitätsbibliothek erworbenen elektronischen Ressourcen, die freien Internet-Ressourcen, die für die Forschung und Lehre an der Humboldt-Universität wichtig sind, und alle Ressourcen, die an der Humboldt-Universität selbst produziert werden, nachgewiesen. Diese Quellen werden – wie in unserem klassischen Katalog – formal (und einheitlich) beschrieben. Zum Beispiel werden Bezeichnung, URL, Ersteller und Betreiber erfasst. Darüber hinaus werden die Ressourcen inhaltlich beschrieben. So wie der Bestand an Monographien und Zeitschriften nach der Regensburger Verbundklassifikation sachlich erschlossen wird, werden auch die Internet-Ressourcen nach einem ähnlichen System klassifiziert. Damit können die für ein Thema wichtigen Internet-Ressourcen an einer Stelle gefunden werden. Die Unterstützung der Mitarbeiterinnen und Mitarbeiter der Universitätsbibliothek bei der »Erwerbung« von Internet-Ressourcen durch die Fachwissenschaftler ist genauso notwendig, wie es bei der klassischen Erwerbung von Monographien und Zeitschriften eingebürgerte Praxis ist.

In der Abb. 1 ist zu sehen, wie der Index der Internet-Ressourcen aussieht:

Neben dem Nachweis der Internet-Ressourcen in einem Katalog soll und kann mit MetaLib die Suche im Internet und in den katalogisierten Internet-Ressourcen unterstützt werden. Wenn eine Internet-Ressource eine Suchmöglichkeit bietet und diese über ein Programm nutzbar ist, soll diese Internet-Ressource so in MetaLib eingebunden werden, dass in dieser Ressource unter der Oberfläche von MetaLib gesucht werden kann. Die Abbildung 2 zeigt, wie das Ergebnis einer Recherche im lokalen Online-Katalog der Universitätsbibliothek, im Katalog der Staatsbibliothek zu Berlin, bei H.Soz.-u. Kult (eine Datenbank des CLIO-Projektes), in den Katalogen der Berliner Öffentlichen Bibliotheken und der australischen Nationalbibliothek dargestellt wird (hier nur die Anzeige der Anzahl der gefundenen Treffer).

Neben der Ermittlung von relevanten Internet-Ressourcen kann auch die Nutzung dieser Internet-Ressourcen unmittelbar an einer Stelle erfolgen. So ist es möglich, gleichzeitig in Datenbanken unterschiedlicher Anbieter zu

suchen, ohne dass Suchbegriffe mehrfach eingegeben werden müssen. Damit dies nicht nur bei Personennamen oder Titelstichwörtern, sondern auch bei inhaltlichem Suchen funktioniert, muss allerdings noch viel Arbeit in die Erstellung von Konkordanzen zwischen den verschiedenen Sacherschließungssystemen der Datenbankanbieter investiert werden.

MetaLib steht zunächst allen eingetragenen Lesern der Universitätsbibliothek uneingeschränkt zur Verfügung. Gäste können je nach Lizenzbedingungen die freien Internet-Ressourcen nutzen oder, wenn sie diese Dienste von Rechnern innerhalb der Humboldt-Universität nutzen, auch die kommerziellen Angebote.

Mit diesem Dienst wird die Universitätsbibliothek ein Portal für die Nutzung des »wissenschaftlichen« Internets aufbauen. Wie es sich für ein solches Portal gehört, kann sich jeder Benutzer seine Umgebung selbst »einrichten«. Neben der Möglichkeit, einmal definierte Anfragen regelmäßig (z. B. täglich, wöchentlich, monatlich, ...) automatisiert abarbeiten zu lassen und sich über Veränderungen per E-Mail informieren zu lassen, kann man sich ein »virtuelles« Buchregal mit den Suchergebnissen speichern. Internet-Ressourcen, die man laufend benötigt, kann man in einer speziellen persönlichen Liste (»Meine Ressourcen-Liste«) verzeichnen und hat damit einen schnellen Zugriff auf diese Quellen.

Abb. 3: SFX-Menü mit Link zum Volltext

Dienste zu den Suchergebnissen

Die Suche nach Informationen im Internet bringt nicht in allen Fällen sofort den Zugriff auf den Volltext. Sehr häufig, insbesondere bei einer Suche in bibliographischen Datenbanken (z. B. in Current Contents oder in Referate- und Rezensions-Datenbanken), findet man Informationen über Dokumente, deren Inhalt dann für die wissenschaftliche Arbeit oder das Studium benötigt wird.

Wenn diese Dokumente im Internet und frei verfügbar sind, ist dies meistens einfach, da mit der Information über die Metadaten (Autor, Titel, ...) meistens ein Link zum entsprechenden Dokument geliefert wird, der angeklickt werden kann und zum Erfolg führt. Mitunter führt ein solcher Link aber zu kostenpflichtigen Angeboten mit allen Problemen, auf die ich weiter oben schon hingewiesen habe. Die Lösung dieser Probleme soll nun mit einem Open-URL-Server erfolgen.

Seit einigen Jahren wird an einem Standard gearbeitet, der inzwischen von immer mehr Datenbank-Anbietern unterstützt wird. Dieser Standard heißt Open-URL [2]. Eine Open-URL dient dazu, innerhalb einer URL Metadaten an einen Open-URL-Server zu übermitteln. Bestandteile einer Open-URL sind ein Identifier, der die Quelle kennzeichnet, die diesen Datensatz liefert, und Metadaten, die das Object selbst beschreiben. Dabei kann es sich um Datenbank-Identifier, um international bekannte Norm-Nummern (ISSN, ISBN, ...) oder um Titel- und Autorennamen handeln.

Auf Grund der Informationen aus der Open-URL kann der entsprechende Open-URL-Server (an der Universitätsbibliothek der Humboldt-Universität ein SFX-Server der Fa. Ex Libris) ermitteln, welche Dienste zu diesem Objekt angeboten werden können. Dazu müssen in einer Datenbank die vertraglichen Bedingungen der Universitätsbibliothek gespeichert werden. Außerdem müssen die Dienste, die von verschiedenen Dienstleistern angeboten werden (Zugriff auf Volltexte, Zugriff auf Abstracts oder Inhaltsverzeichnisse, Dokumentlieferung in Kopie oder als physische Einheit, Kauf, ...), und die, die von anderen Quellen benötigt werden (eine Referate-Datenbank, die Metadaten zu den rezensierten Artikeln gespeichert hat, »fordert« als Dienst z. B. den Zugriff auf den rezensierten Artikel oder zu anderen Artikeln des gleichen Autors an), definiert werden. Aus

diesen Informationen kann der Open-URL-Server für jeden Treffer ein »maßgeschneidertes« Dienstleistungsmenü erzeugen.

Abb. 3 zeigt ein Menü, das u. a. den Zugriff auf den Volltext und zum lokalen OPAC enthält (das Medium befindet sich in der Universitätsbibliothek der Humboldt-Universität), während in Abb. 4, da es keinen Volltext im Online-Zugriff gibt, nur ein Dokumentlieferdienst angeboten wird.

Zurzeit sind in unserem SFX-Server die Dienste für die wichtigsten elektronischen Zeitschriften und eine Reihe von Datenbanken konfiguriert.

Wie kann nun der SFX-Server eingesetzt werden? Zunächst einmal ist MetaLib so konfiguriert, dass für alle Ressourcen, in denen mit Hilfe von MetaLib recherchiert werden kann, Open-URL's erzeugt werden. Damit ist es möglich, egal in welcher Datenbank eine Recherche begonnen wurde, mit einem Klick auf den SFX-Link zu ermitteln, ob das entsprechende Dokument Online im Volltext erreicht werden kann, ob es sich im Bestand der Universitätsbibliothek befindet oder ob es kostengünstig über subito.Dokumente aus Bibliotheken e. V. bzw. die Fernleihe beschafft werden kann.

Darüber hinaus wird der SFX-Server aber noch in zwei anderen Fällen wirksam. Die Universitätsbibliothek wird mit den Datenbankanbietern, die Open-URL-enabled[2] sind, vereinbaren, dass für die Mitglieder der Humboldt-Universität bei Recherchen in diesen Datenbanken entsprechende Open-URL's erzeugt werden, die auf unseren SFX-Server verweisen.

Als zweite Variante sei der »Zitation-Linker« genannt, der auf speziellen WWW-Seiten (*http://www.ub.hu-berlin.de/ citation-linker*) die Möglichkeit anbietet, direkt bibliographische Daten (Autor, Titel, ISBN, ISSN, ...) einzugeben und die Dienste zu ermitteln, die die Universitätsbibliothek zu diesen Daten anbietet. Diesen Dienst sollte man dann benutzen, wenn man in einer Arbeit Referenzen gefunden hat. Dieser Dienst ersetzt natürlich nicht die anderen Such- und Recherchemöglichkeiten innerhalb der Universitätsbibliothek und im Internet.

2 Open-URL-enabled bezeichnet solche Datenbanken, die beim Aufbau einer Treffermenge eine Open-URL erzeugen (können).

Abb. 4: SFX-Menü mit Link auf Abstract und Inhaltsverzeichnis

MetaLib und SFX als Dienstleistung der Universitätsbibliothek für die Institute der Humboldt-Universität

Mit den oben beschriebenen Möglichkeiten bietet die Universitätsbibliothek ihren Lesern neue Möglichkeiten und eine einfachere Navigation im Internet. Wir wollen aber noch einen Schritt weiter gehen. Wir wollen diese Produkte, die von uns lizenziert und gewartet werden, den Forschungsgruppen, Instituten und Einrichtungen der Humboldt-Universität zur unmittelbaren Nutzung anbieten. Was soll das heißen?

Wenn Forschungsgruppen Gateway's für das eigene Fachgebiet aufbauen wollen, in denen Suchmöglichkeiten über unterschiedliche Datenbanken integriert werden sollen, und wenn von Suchergebnissen unterschiedliche Dienste angeboten werden sollen, kann dies mit Hilfe von MetaLib und SFX erfolgen.

Mit dem Projekt »CLIO-online« [3] haben wir die ersten Schritte in diese Richtung getan.

Literatur

[1] http://www.exlibris.co.il
[2] http://www.sfxit.com
[3] http://www.clio-online.de; DFG-Projekt am Institut für Geschichtswissenschaften