

Sonderdruck
Nr. 14

8/2005

Information

WISSENSCHAFT & PRAXIS



Barrierefreiheit im Internet

Zielgruppen

Rechtliche Grundlagen

Barrierefreies Webdesign

Qualitätsmanagement für Websites

Sprachliche Hürden

Barrierefreie PDF-Dokumente

BITV-Test zur Barrierefreiheit

BIENE-Award

**Praxisbeispiel zur Einführung
von Barrierefreiheit**

**Realisierung barrierefreier
Web-Angebote**

Jahresregister 2005



www.bookeye.com



Das 18oT-Projekt in Köln oder wie verarbeite ich 180.000 Bücher in vier Monaten

Eine erfolgreiche Kooperation des hbz, der USB Köln und der ZB MED

Astrid Großgarten, Bonn

Versetzen wir uns in die Lage eines Studenten der Sozialwissenschaften, der Literatur für eine Seminararbeit zum Thema „Die Rolle der Frau im ländlichen Japan an der Schwelle zum 21. Jahrhundert“ sucht oder in die einer Dermatologin, die zu „Behandlungsmethoden bei Rosacea“ forscht. Die Katalogrecherche wird neben Standardwerken auch Treffer zu Monographien aufweisen. Aber werden die Suchenden auch fündig in Aufsatzsammlungen wie „Japanese woman working“ oder „Naturheilverfahren für Hauterkrankungen“? Diese Frage lässt sich nur mit einem Blick in das Inhaltsverzeichnis beantworten. Befinden sich die Recherchierenden gerade in der Bibliothek, ist dies unproblematisch. Anders sieht es jedoch aus, wenn online recherchiert wird. Was tun? In die Bibliothek fahren und dort nachsehen? Das Buch auf Verdacht per Fernleihe ordern? Eine Erweiterung der Kataloginformation um die Inhaltsübersicht wäre also wünschenswert. Solche Wünsche werden bald Wirklichkeit - am schnellsten für Mediziner, Betriebswirt- und Sozialwissenschaftler in wissenschaftlichen Bibliotheken in Köln.

Unter dem Stichwort *Catalogue Enrichment* läuft derzeit in Köln ein Projekt, in dessen Verlauf die Inhaltsverzeichnisse von 180.000 Büchern gescannt, mit einer Texterkennung als Volltext aufgearbeitet und schließlich in die verschiedenen Katalogsysteme eingespeist werden. Das Projekt startete am 1. September 2005 und soll vor Weihnachten abgeschlossen werden. Noch läuft das Unterfangen unter dem Arbeitstitel 18oT-Projekt (für 180.000 Bücher). Einen endgültigen Titel wird das Projekt in einer späteren Phase bekommen, denn, wie Frau Gitmans von der Projektkoordination des hbz (Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen) betont, es werde über die Fortsetzung des *Catalogue Enrichments* mit weiteren Partnerbibliotheken nachgedacht, so dass die Menge dann 180.000 Medien überschreite und ein neuer Titel gefunden werden müsse.

Wie kam es zu einem derartigen Projekt, das in seiner Größenordnung bislang einmalig in der deutschen Bibliothekslandschaft ist? Seit einiger Zeit steht das Thema *Catalogue Enrichment* im Raum, wie Tatjana Mrowka, die Marketingleiterin des Projektträgers hbz sagt. Auf verschiedenen Fachtagungen wurde das Thema immer wieder beleuchtet – insbesondere unter dem Aspekt, wie durch zusätzliche Inhaltsinformationen ein Mehrwert für den Katalog (OPAC) geschaffen werden kann, um die Literatursuche zielgerichteter und erfolgreicher zu gestalten. Natürlich sehen sich Bibliotheken bei der Allgegenwart von Amazon und Google und verwöhnten Internetnutzern auch stark gefordert, das Feld der Innovation nicht allein kommerziellen Anbietern zu überlassen.

Auf der Konferenz des hbz-Verbundes zu Beginn des Jahres 2005 wurde die Diskussion um *Catalogue Enrichment* erneut aufgegriffen, und das hbz beschloss, ein Projekt in größerem Stile anzugehen. Unterstützt wird das Projekt vom Ministerium für Innovation, Wissenschaft, Forschung und Technologie des Landes Nordrhein-Westfalen.

Tabelle 1: Einige Kennzahlen aus dem vielfältigen Produktspektrum des hbz

Verbunddatenbank	13.750.000 Titeleintragungen 30.200.000 Bestandsnachweise 246 Teilnehmerbibliotheken
Die Digitale Bibliothek	330 eingebundene Datenbanken 170 wissenschaftliche und öffentliche Bibliotheken
Online-Fernleihe	knapp 300 Bibliotheken

Das hbz ist aufgrund seiner Zielsetzung prädestiniert für die Federführung in einem solchen Pilotprojekt, da es als Dienstleistungs- und Entwicklungseinrichtung innovative Tendenzen aufgreift. Mit einem breitgefächerten Produktspektrum ist das hbz bundesweit aktiver Partner von Bibliotheken, die insgesamt 2,5 Millionen Kunden mit Literatur und Informationen versorgen.



Freihandausleihe ZB MED Köln

Das hbz als Projektkoordinator stimmte in Kooperation mit der Universitäts- und Stadtbibliothek Köln (USB Köln) und der Deutschen Zentralbibliothek für Medizin (ZB MED) die Eckdaten ab. Aus dem Bereich der ZB MED werden 60.000 Monographien aus den Zugängen der letzten fünf Jahre bearbeitet, aus der USB Köln 120.000 Titel aus den Erwerbungen der letzten 15 Jahre des Fachbereiches Wirtschafts- und Sozialwissenschaften. Der unterschiedliche Zeitanatz bei dieser retrospektiven Kataloganreicherung ist der Tatsache geschuldet, dass die Halbwertszeit von medizinischem Wissen eine wesentlich kürzere ist als auf dem Gebiet der Wirtschaft- und Sozialwissenschaften.

Beide teilnehmenden Bibliotheken haben eine herausgehobene Stellung. So ist die ZB MED die zweitgrößte medizinische Fachbibliothek der Welt, nach Nutzerzahlen gar die größte. Die USB Köln spricht mit mehreren DFG-Sondersammelgebieten, umfangreichen Spezialbeständen und dem bedeutendsten Altbestand in NRW auch überregionale und außeruniversitäre Kunden an. Für das Projekt wurden Titel aus

dem betriebs- und sozialwissenschaftlichen Bestand gewählt, der etwa ein Drittel des Gesamtbestandes ausmacht. Beide Fachbereiche decken neben der deutschen und angloamerikanischen Forschungsliteratur auch weitestgehend die relevanten Titel aus dem gesamten europäischen Sprachraum ab.

Tabelle 2: Kennzahlen der beteiligten Bibliotheken (Stand 2004)

	USB Köln	ZB MED
Bestand an Büchern und Zeitschriftenbänden	3,6 Millionen	1,3 Millionen
Laufende Zeitschriftentitel	ca. 10.000	ca. 8.000
Nutzer im Jahr	über 45.000	14.815
Ausleihvorgänge am Ort	1.170.000	147.000
Gebende Fernleihen	ca. 60.000	
Dokumentenlieferungen und Fernleihen		610.000

Die wissenschaftliche Relevanz der Fachliteratur aus den ausgewählten Pilotbibliotheken und die Tatsache, dass im Verbundkatalog des hzb die Daten von 246 Bibliotheken zusammenfließen, macht deutlich, wie sinnvoll es ist, angereicherte Titeldaten in einen solch großen und stark frequentierten Datenpool wie den hzb-Medienserver einzustellen. Darin stehen den Verbundteilnehmern gegenwärtig 13 Millionen Titeldaten mit etwa 30 Millionen Exemplardaten zur Verfügung.

Rechnet man die für das Projekt ausgewählte Bücherzahl von 180.000 auf zu scannende Seiten hoch, liegt die Gesamtseitenzahl bei 720.000, von denen rund 240.000 Seiten auf Medizin und rund 480.000 Seiten auf Wirtschafts- und Sozialwissenschaften entfallen. Umgerechnet auf vier Monate entspricht dies also etwa 2000 Büchern oder 7000 Seiten am Tag, die zu verarbeiten sind.

Für das hzb und die teilnehmenden Bibliotheken stand von Beginn an fest, dass eine

derartige Größenordnung nicht im normalen Bibliotheksalltag vom eigenen Personal zusätzlich bewältigt werden kann.

Gesucht wurde nun ein Dienstleister, der die zeitlichen, organisatorischen und finanziellen Vorgaben erfüllen konnte. Diese sahen vor, dass für das Projekt keine zusätz-

liche Hard- und Software angeschafft werden durfte, auf der vorhandenen Infrastruktur (Medea3-Umfeld der hzb-Verbundbibliotheken und MyBib-Server der USB Köln) aufgesattelt werden musste, keine erneute Mediendatenerfassung stattfinden sollte und natürlich der Bibliotheksbetrieb im Ablauf nicht behindert werden durfte.

Die Firma ImageWare Components¹ aus Bonn, bekannter Hersteller von Bookeye®-Buchscannern und MyBib-Liefersystemen, konnte sich schließlich als Anbieter mit der überzeugendsten Lösung durchsetzen.

Gemeinsam wurde die Projektorganisation verabredet:

- die beteiligten Bibliotheken stellen Netzwerkverbindungen und Arbeitsräume für den Dienstleister und ermöglichen dem Dienstleistungspersonal Zutritt zu Freihandausleihe und Magazinen
- der Dienstleister stellt die technische Ausstattung
- der Dienstleister stellt einen Server für die Auftragsbearbeitung, der wiederum über eine Schnittstelle mit dem hzb-Server verbunden wird
- das hzb konfiguriert zusammen mit dem Dienstleister den Server
- in beiden beteiligten Bibliotheken gibt es feste Ansprechpartner für technische Probleme, Qualitätssicherung und fachliche Fragen
- gemeinsam werden die Qualitätsstandards festgelegt und deren Einhaltung überprüft.

Für ImageWare Components bedeutete der Projektauftrag eine große Herausforderung, betrat man doch durch die Komplexität der Anforderungen auch teilweise technisches und unternehmerisches Neuland. Geschäftsführer Rolf Rasche sah weit mehr Chancen als Risiken, „denn die Möglichkeit, MyBib einem solchen Belastungstest vor der Haustüre zu unterziehen, bekomme ich nicht oft geboten. Außerdem können wir so die Stabilität und Qualität unserer Systeme beweisen“.

Um Überraschungen bei Projektstart zu vermeiden, war eine Pilotphase im Juli und August vorgeschaltet. Zunächst wurden jeweils 500 Bände pro teilnehmender Bibliothek verarbeitet. Die Ergebnisse wurden von den Ansprechpartnern der Bibliotheken und des hzb hinsichtlich Scanqualität und Genauigkeit der Texterkennung überprüft. Das Ergebnis war mehr als zufriedenstellend, ebenso wie der optimierte MyBib-edoc-Server, über den die Auftragsverwaltung, -steuerung, -verfolgung und Betriebsdatenerfassung lückenlos abgewickelt wird. „Für das hzb steht die Qualitätssicherung an erster Stelle. Wären die Testergebnisse nicht zufriedenstellend gewesen, hätten wir das Projekt eingestellt.“, so Hans Ollig, kommissarischer Leiter des hzb.

Nach der erfolgreichen Testphase ging das Projekt am 1. September 2005 in den Produktionsbetrieb und läuft seitdem zur vollsten Zufriedenheit aller Beteiligten. Da die Arbeitsergebnisse jederzeit von jedem Projektmitglied per Web-Zugriff auf das MyBib System geprüft werden können, ist die Projektgruppe stets auf dem Laufenden, und man trifft sich nur einmal im Monat. Auf der Tagesordnung steht dabei die Bewertung der Arbeitsqualität. Zur Diskussion kommen dann auch Sonderfälle, für die zeitnah verbindliche Vorgaben und Lösungen für den Produktionsbetrieb gefunden werden müssen.

Wie sieht es nun in den teilnehmenden Bibliotheken aus? In der USB Köln wurden acht und in der ZB MED vier Scanstationen aufgebaut. Die Ausrüstung pro Arbeitsplatz sind neben Mobiliar und Bücherwagen jeweils ein Bookeye®-GS400 mit ergonomischem Scanpad und Barcodepistole. Ansgret Johann von ImageWare, verantwortlich für die Produktion, erläutert den Ablauf: „Die Mitarbeiter holen die Bücher an den Scanplatz. Dort wird jedes Buch zuerst mit der Barcodepistole registriert. In den folgenden Arbeitsschritten werden die Seiten des Inhaltsverzeichnisses gescannt und um irrelevante Informationen bereinigt. Danach wird die Texterkennung durchgeführt. Jeder Mitarbeiter ist angehalten, die Stimmigkeit des Ergebnisses zu überprüfen“. Sollten Fehler übersehen worden sein, fallen sie später in dem mehrstufigen Qualitätssicherungsverfahren auf. Dann bekommen die Aufträge im System einen Reklamationsvermerk und erscheinen auf späteren Auftragslisten zur Nachbearbeitung.

Die Mitarbeiter, die die Scanarbeiten ausführen, sehen nur die Benutzeroberflächen der verwendeten Scansoftware BCS-2®. Das



Scanstation in der USB Köln

¹ ImageWare Components GmbH, Am Hofgarten 20, 53113 Bonn, Telefon: (02 28) 96 98 5-0, Fax: (02 28) 96 98 5-84, E-Mail: info@imageware.de, www.imageware.de



Erfassung in der ZB Med Köln

recht aufwändige MyBib-System im Hintergrund, das den gesamten Workflow steuert und die lückenlose Auftragsverfolgung von den lokalen Bibliothekssystemen zum hbz-Server erst möglich macht, ist nur den jeweiligen Projektbeauftragten zugänglich. Über MyBib werden von Frau Johann sogenannte Buchhollisten erzeugt und an die Scanoperatoren verteilt. Diese Listen sind Auftragszettel und geben den Mitarbeitern die zu bearbeitenden Bücher vor. Jedes Buch ist über Signatur, Mediennummer und Titel auf der Liste ausgewiesen. Den Mediennummern kommt eine Schlüsselrolle in dem komplexen Datengefüge zu. Sie ist die eindeutige Identifizierung für ein Buch und in Form eines Barcodeetiketts auf jedem Medium aufgebracht. Der Barcode identifiziert das Buch gegenüber MyBib und dies stellt wiederum eine Verknüpfung zur Verbund-ID des hbz her.

Gab es nun im Echtbetrieb unvorhergesehene Schwierigkeiten oder Fragestellungen? Nicht wirklich, so unisono die Meinung aller Beteiligten, allenfalls tauchten Sonderfälle auf, über deren verbindliche Handhabung man sich in den monatlichen Projektsitzungen oder über eine Mailingliste verständigt. Was sind denn nun Sonderfälle? „Kritzeleien im Inhaltsverzeichnis, mehrsprachige Verzeichnisse und solche mit Formeln oder arabischen und chinesischen Schriftzeichen werden von der Texterkennung nicht oder nur fehlerhaft erkannt.“, so Frau Johann.

Da das Projekt zeitlich voll im Plan liegt und reibungslos läuft, schauen alle Beteiligten schon weiter in die Zukunft und prüfen die Möglichkeit einer Fortführung, bei der man dann die Kataloganreicherung nicht nur retrospektiv, sondern auch für Neuzugänge betreiben könnte. Um möglichst 80 Prozent der Neuzugänge im Verbundbereich abzudecken, ist geplant, das Projekt in einer zweiten Phase zunächst mit fünf bis sechs Bibliotheken fortzusetzen.

Bevor dies soweit ist, steht aber noch die Vollendung des 180T-Projekts an und damit auch der konkrete künftige Nutzen aus der Maßnahme.

Das hbz wird zum Jahreswechsel die gewonnenen Daten in seinen Medienserver übernehmen. Die Umsetzung des durch Texterkennung generierten Volltextes er-

laubt dabei eine Indexierung der Daten über die im hbz verwendete Suchmaschine, die auf der Technologie FAST beruht. Die angereicherten Daten werden für alle Verbundteilnehmer nutzbar sein und doppelte Arbeit wird vermieden.

Jeder Bibliothekskunde profitiert davon, dass die Einträge von 180.000 Monographien in den Katalogen um die Inhaltsverzeichnisse erweitert werden. Diese werden für die Nutzer dann als digitales Bild einsehbar sein. So wird die Literaturrecherche für die Orts- und Fernleihe erheblich verbessert und erleichtert. Als Nebeneffekt werden nun vielleicht auch Dokumentenlieferaufträge für Beiträge erteilt, die sonst weiter unbeachtet in Sammelbänden im Regal verblieben wären.

Für Dr. Christiane Süverkrüp und Claudia Dembek, die das Projekt für ihre Bibliotheken inhaltlich betreuen, liegt der Vorteil vor allem in der verbesserten Arbeitsökonomie der Bibliothekskunden, denn die ergänzten Katalogdaten werden Fehlausleihen mindern und die Nutzungsfrequenz der Titel erhöhen. „Durch das Catalogue Enrichment und die Online-Verfügbarkeit der Daten werden die vorhandenen Bestände zudem besser publik gemacht und der Nutzungsservice gerade auch für die überregionalen Kundenkreise erweitert.“, so Frau Dembek. Nach der Erfahrung von Frau Dr. Süverkrüp hat besonders die Online-Recherchierbarkeit von Medien eindeutige Auswirkungen auf deren Nutzungshäufigkeit. Und was kann sich eine Bibliothek für ihr erhebliches Investment in Bestandsaufbau und -pflege mehr wünschen, als dass ihre Medien zielorientiert, intensiv und

häufig genutzt werden? Und vielleicht kehrt der ein oder andere Nutzer dann zukünftig von Google zurück und recherchiert bei der nächsten Gelegenheit lieber wieder in den Katalogen der ZB MED und der USB Köln.

Projekt, Optische Zeichenerkennung, Katalog, Monographie, Literaturdokumentation, Stadtbibliothek Köln, Deutsche Zentralbibliothek für Medizin, Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen, Catalogue Enrichment

DIE AUTORIN

Astrid Großgarten



M.A., studierte Spanisch, Geschichte und Sozialwissenschaften. Nach dem Studium arbeitete sie mehrere Jahre im Bereich Marketing und Vertrieb mittelständischer Unternehmen. Seit 2004 als Verlegerin für die Edition historischer Bücher auf CD tätig.

Patrimonium Transcriptum
Verlags GmbH
Oppenhoffstraße 16
53111 Bonn
E-Mail: grossgarten@ptverlag.de
Telefon: (02 28) 3772 685



Hochschulbibliothekszentrum des
Landes Nordrhein-Westfalen
Jülicher Straße 6
D - 50674 Köln
Telefon: 0221 / 400 75 - 0
Telefax: 0221 / 400 75 - 180
E-Mail: info-hbz@hbz-nrw.de
<http://www.hbz-nrw.de>



ImageWare
Components GmbH
Am Hofgarten 20
D - 53113 Bonn
Telefon: 0049- 228/96985 - 0
Telefax: 0049- 228/96985 - 84
E-Mail: info@imageware.de
<http://www.imageware.de>