

Andreas Bohne/Jörg Schwiemann

Entwurf und Realisierung einer Web-basierten Verknüpfung einer Literaturlatenbank mit einem Bestell- und Recherchesystem



Der erste Teil des Artikel beschreibt das Konzept und die Realisierung der Verknüpfung einer Web-basierten Literaturlatenbank mit einem Bestell- und Recherchesystem (Subito) für Artikel. Im zweiten Teil wird eine für das Deutsche Krebsforschungszentrum Heidelberg entwickelte Installation eines Internet-basierten Fernleihsystems vorgestellt.

Concept and implementation of a literature database hyperlinked with an order and retrieval system

The first part of the article describes concept and implementation of a literature database hyperlinked with the retrieval and order system Subito. The second part reports an Internet-based interlibrary loan request system installed at the German Cancer Research Center Heidelberg.

Conception et réalisation d'un hyperlien entre une base de données bibliographique et un système de recherche et de commande

Cet article décrit la conception et la réalisation d'un hyperlien entre une base de donnée bibliographique accessible par la toile et le système Subito pour la recherche et la commande de monographies et d'articles. La deuxième partie de l'article rapporte un système réalisé au centre de la recherche du cancer à Heidelberg construit pour le prêt interbibliothèque.

Einleitung

Die ganze Welt ist vernetzt – aber dort, wo man es manchmal braucht, werden Daten nicht weitergereicht. Auch wenn inzwischen die vierte Phase des Internets, nämlich die Verknüpfung von Web Application Servern, eingeläutet worden ist, so ist dies momentan ein Bereich, der nur das große Internet-Business betrifft. Geplant ist, dass man z.B. ein Buch bei einem Online-Buchladen bestellen kann, und wenn die Sendung ausgeliefert wird, den Weg der Sendung online verfolgen kann, um eventuell nachzuschauen, wo die Buchsendung gerade aufgehalten wird. Das Kaufen und das online-Verfolgen des Buches ist im einzelnen schon realisiert – das Zusammenspiel wird gerade erst entwickelt. Bisher tut man sich schwer, Daten von einem Webformular in ein anderes zu übertragen, besonders dann, wenn beide Formulare auf einem fremden Server liegen. Und doch zeigt die Praxis allzu oft, wie wichtig es ist, Schnittstellen zu definieren und Webanwendungen miteinander zu verknüpfen. Auch scheinen die Entwickler von Web-Browsern noch keinen Vorteil in einem regelbasierten Formularzwischenspeicher zu sehen. Sicherlich wäre es schön, wenn man als Regel bei einem Webbrowser definieren könnte: ‚Füge bei einer PubMed-Abfrage die Zeile 3 aus der Ergebnisseite beim Aufruf der Seite <<http://www.subito-doc.de>> in das erste Formulartextfeld ein.‘ Technisch wäre die Realisierung möglich, praktisch scheint der Bedarf nicht groß genug zu sein, um Zeit für die Realisierung zu investieren. Dabei wird einem die Sache angesichts der zunehmenden Verbreitung von XML geradezu in den Schoß gelegt.

Anwendung

Konkret ging es bei unserem Problem darum, dass das Deutsche Krebsforschungszentrum Heidelberg (DKFZ) eine Literaturlatenbank¹ (Sweet-DB) pflegt, in der relevante Struktur- und Literaturdaten von Kohlenhydraten

gespeichert sind, und diese Datenbank sollte mit dem Bestell- und Recherchesystem von Subito online verbunden werden, so dass es möglich ist, einen gewünschten Artikel online zu bestellen. Hervorgegangen ist die Sweet-DB aus einem in Amerika initiierten Projekt, der CarbBank², das aus Mangel an finanziellen Mitteln eingestellt worden ist. Dabei wurden in der aktiven Phase des Projektes die Daten von Hand erfasst und abgelegt. So wurden Stück für Stück 49 000 Einträge zusammengeführt. Dabei umfasst ein Eintrag z.B. folgende Werte:

```
; start of record
; db=ccsd29
; Record #=1
CC: CCSD:42034
AU: Albersheim P; Darvill A; Augur C; Cheong JJ;
Eberhard S; Hahn MG; Marfa V; Mohnen D;
O'Neill MA; Spiro MD; York WS
TI: Oligosaccharins: Oligosaccharide regulatory
molecules
CT: Acc Chem Res (1992) 25: 77-83
SC: 1
BS: (GS) Phytophthora megasperma, (OT) cell wall,
(*) f.sp. glycinea
SB: Westra B
DA: 06-12-1995
FC: 5bc58ff3
SI: CBank:15402
-----
structure:
  b-D-Glcp- (1-6) +
      |
      b-D-Glcp- (1-6) -b-D-Glcp- (1-6) +
      |
  b-D-Glcp- (1-3) +
      |
      b-D-Glcp- (1-6) -D-Glc
      |
      b-D-Glcp- (1-3) +
=====end of record
```

- 1 SWEET-DB: <<http://www.dkfz.de/spec2/sweetdb/>>.
- 2 CarbBank: Die originalen Seiten existieren nicht mehr. Angesiedelt waren sie am CCRC (<<http://www.ccr.c.uga.edu/>>).

Abb. 1 Beispiel Sweet-DB

Man sieht, dass neben anderen Daten, wie Analysemethode und biologische Bedeutung, auch die genauen Literaturdaten wie Autor, Titel, Journal, Jahr, Bandnummer und Seitenangaben abgelegt sind. Ist man nun an dem Artikel interessiert und hat man das Pech, dass der Artikel in einer Zeitschrift veröffentlicht worden ist, die nicht im aktuellen Bestand der Bibliothek vor Ort geführt wird, muss man den Weg über die Fernleihe gehen.

Eine Möglichkeit dabei ist der Anbieter Subito³. Die Firma Subito ist aus einem BMBF-Projekt hervorgegangen und koordiniert Fernleihaufträge. Dabei stellt Subito eine WWW-Schnittstelle zur Verfügung, über die ein Artikel bestellt werden kann. Neben einem Gastzugang gibt es auch die Möglichkeit, sich als registrierter Benutzer anzumelden.

Der eigentliche Bestellvorgang ist mehrstufig aufgebaut. Zuerst meldet sich der Benutzer bei dem System an. Dabei wird ihm intern eine Session-ID zugeordnet. Über diese Session-ID wird zum Beispiel geprüft, ob der Benutzer einen Bestellvorgang innerhalb einer bestimmten Zeit abgeschlossen hat. Dieses soll Missbrauch verhindern, nämlich, dass ein Benutzer sich bei dem System anmeldet, seine Bestellung vornimmt, dann aber vergisst, sich abzumelden und den Rechner verlässt.

Nach dem Anmelden sucht der Benutzer in der momentan realisierten Oberfläche zuerst die Zeitschrift, aus der er den Artikel bestellen möchte. Werden bei der Suche nach der Zeitschrift mehrere Treffer gefunden, sei es, weil die Zeitschrift umbenannt worden ist, oder weil mit Trunkierungszeichen gearbeitet worden ist, so wählt der Benutzer auf der zweiten Seite die Zeitschrift aus, die er möchte. Auf der folgenden Seite werden die Bibliotheken angezeigt, die die ausgewählte Zeitschrift in ihrem Bestand haben. Der Benutzer wählt dort die Bibliothek aus, bei der er die Zeitschrift bestellen möchte. Hiernach kommt er zu der Seite, wo er Seitenzahl, Bandnummer und Jahr eingibt. Die Angabe von Autor und Titel sind optional. Nach diesem Schritt geht die Bestellung in die letzte Runde, wo der Benutzer Anschrift und Zahlungsmodalitäten eingibt, falls er nicht registrierter Benutzer ist.

Bisher gab es jedoch zwei Webseiten, auf denen der Benutzer Daten eingeben musste, die auf dem System des DKFZ schon in elektronischer Form vorhanden sind. Da es sich um zwei verschiedene Web-Server handelt (DKFZ-Datenbank und das Subito-Recherche- und Bestellsystem), war eine Verknüpfung nicht ohne weiteres möglich. Auch die durch das Sitzungs-Konzept bedingte mehrstufige Vorgehensweise bei der Bestel-

lung auf dem Subito-Rechner ließ eine ad hoc-Lösung, wie das lokale Ausfüllen eines Webformulars und anschließende Übermittlung an den Subito-Rechner, nicht zu. Subito selbst arbeitet zwar an einer PreOrder-Lösung (einstufiges Bestellverfahren), die jedoch nicht vor Ende des Jahres 2001 verfügbar sein

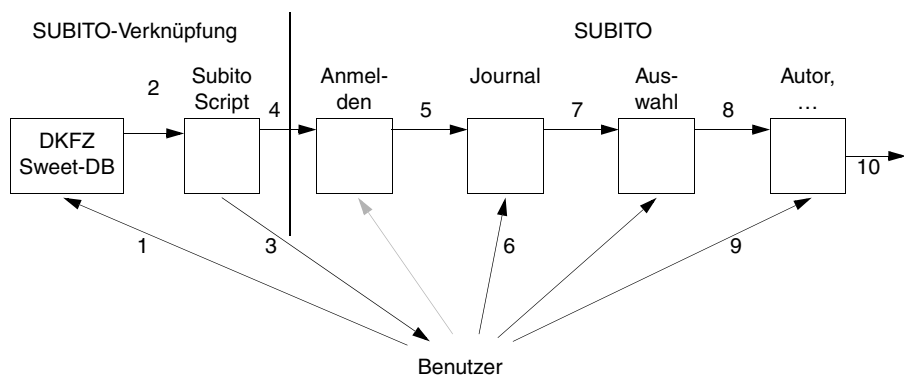
wird. Um sich dennoch in das System einklinken zu können, war eine der wichtigsten Nebenbedingungen, dass das Subito-System so wenig wie möglich verändert werden sollte. Ferner sollte eine Lösung gefunden werden, die auf allen Standard-Web-Browsern läuft.

Zu diesem Zweck wurde der Weg über Cookies⁴ gewählt. Cookies, oft verschrien als gefährlich, sind lediglich Daten, die in einer Datei auf dem Rechner des Benutzers lokal gespeichert werden. Diese Datei kann man jederzeit anzeigen lassen oder löschen. Cookies können nur an dem Rechner gelesen werden, von dem sie auch gesetzt worden sind. Es handelt sich dabei lediglich um Textdaten. Cookies sind keine ausführbaren Programme und sie können auch keine Kreditkartennummern ausspähen und versenden. Da Cookies jedoch immer in einem Atemzug mit Datenschutz und Privatsphäre genannt werden, soll an dieser Stelle kurz erläutert werden, wo die Gefahr liegt. Angenommen, Sie haben bei einem Online-Buchhändler ein Buch erworben. Zu diesem Zweck haben Sie das Buch ausgewählt und sind dann an die virtuelle Kasse gegangen und haben dort ihren Namen und Ihre Adresse hinterlassen. Das Online-Shop-System hat nun ein Cookie gesetzt. In diesem Cookie steht für gewöhnlich eine Zahl. Diese Zahl weist Sie in dem Online-Shop als Benutzer XY aus. Angenommen, Sie stöbern eine Woche später in demselben Online-Shop, so kann das System Sie aufgrund der Nummer in dem Cookie identifizieren und die Suche Ihrem Benutzerprofil hinzufügen – ohne dass Sie sich dem System als Benutzer XY ausgewiesen haben.

In dem vorgestellten System werden die Daten Autor, Titel, Bandnummer, Jahr, Seite und Journal in dem Cookie abgelegt. Das Cookie hat eine Lebensdauer von 10 Minuten und wird von dem Subito-Rechner gesetzt, da auch der Subito-Rechner während der einzelnen Phasen des Bestellens die Cookies liest. Bei einem Benutzer, der nicht über die DKFZ-Zuckerdatenbank zu Subito verzweigt, werden auch keine Cookies gesetzt. Ein Link von der Zuckerdatenbank setzt erst via Cookies die Literaturdaten über den Subito-Rechner und geht dann zu der normalen Bestellseite weiter. Auf der ersten Subito-Seite, auf der die Zeitschrift eingegeben werden muss, überprüft die Seite via JavaScript, ob der Benutzer gesetzte Literaturdaten in seinem Cookie hat. Wenn ja, so werden diese automatisch eingetragen. Der Be-

3 Subito: <<http://www.subito-doc.de>>.

4 Cookie-Informationen: <<http://www.cookiecentral.com>>.



- 1) Der Benutzer wählt eine Literaturstelle in der Glykodatenbank aus.
- 2,3) Ein CGI-Script setzt auf dem Rechner des Benutzers ein Cookie mit den Daten: Autor, Journal, Jahr, etc.
- 4) Der Benutzer beginnt einen normalen Bestellvorgang.
- 5,6) Bei der Seite mit der Auswahl der Zeitschrift wird beim Laden überprüft, ob gültige Literaturangaben in einem Cookie von Subito vorhanden sind. Wenn ja, dann werden diese automatisch eingesetzt.
- 7) Der Bestellvorgang wird weitergeführt.
- 8,9) Bei der Seite mit der Eingabe des Zeitschriftenbandes, Jahr und Seitenangaben wird beim Laden überprüft, ob gültige Literaturangaben in einem Cookie von Subito vorhanden sind. Wenn ja, dann werden diese automatisch eingesetzt.
- 10) Der Bestellvorgang kann weitergeführt und abgeschlossen werden.

Abb. 1 Verknüpfung DKFZ-Datenbank und Subito

nutzer macht nun ganz normal im Bestellvorgang weiter. Auf der Seite, wo er normalerweise Bandnummer und so weiter eingeben muss, wird erneut überprüft, ob Daten vorhanden sind, und wenn ja, werden auch diese automatisch eingetragen. Das JavaScript konnte auf der Subito-Seite problemlos in die HTML-Schablonen der dynamischen Web-Seiten eingebunden werden. Als Nebenprodukt dieses Projektes entstand eine Web-Schnittstelle für die Zentralbibliothek des DKFZ. Bisher wurden Fernleihscheine auf Papier ausgefüllt, gesammelt und von einer Angestellten in den Computer eingegeben. Dieses stellt einen klassischen Medienbruch dar. Teilweise haben Bibliotheksbenutzer in PubMed recherchiert, das Ergebnis ausgedruckt, die einzelnen Artikelangaben ausgeschnitten und auf einzelne Fernleihscheine aufgeklebt. Alle Fernleihen des DKFZ werden über die Bibliothek abgewickelt. Somit tritt nicht der nor-

male Benutzer als Kunde von Subito auf, sondern nur die Bibliothek, welche auch die Abrechnung und die Weiterleitung an die Benutzer übernimmt. Mit der neuen Web-Oberfläche kann der Benutzer eine Fernleihe direkt von seinem Arbeitsplatz im DKFZ tätigen. Das Webformular ist nur von den internen Rechnern aus zugänglich.

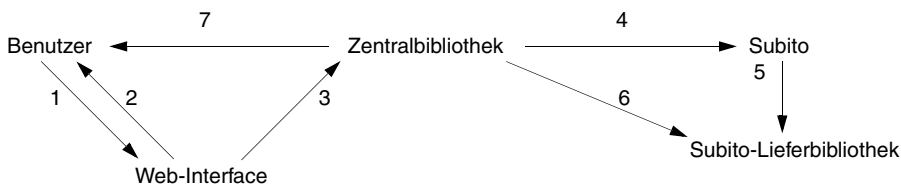
Durch diesen interaktiven Ansatz ist es möglich, fehlende Angaben schon bei der Eingabe zu überprüfen, denn es ist schlecht, eine Fernleihe durchzuführen, wenn die Jahresangabe fehlt. Hat der Benutzer neben den Literaturangaben auch seinen Namen, seine Email-Adresse und seine Kostenstelle eingegeben, so wird über ein PHP4-Script aus den Eingaben eine Mail erzeugt. Während der Bestellung wird noch das Datum automatisch eingefügt. Der Besteller bekommt eine Mail mit einer Bestätigung, in der ihm zur Kontrolle mitgeteilt wird, welche Fernleihe er gerade getätigt hat. Eine zweite Mail geht an die Fernleihabteilung der Zentralbibliothek. Diese Mail wird dort automatisch in ein entsprechendes Mail-Eingangsverzeichnis einsortiert, aus dem sie zur Bearbeitung dann wieder abgerufen werden kann. Die Mail an die Fernleihabteilung besteht im Gegensatz zu der Mail an den Benutzer aus zwei Teilen. Zum einen erhält sie natürlich die Angaben des Benutzers in ASCII-Form zum Ausdrucken und Ablegen, und zum anderen enthält diese Mail ein Web-Formular in HTML-Code. Dieses wurde über die Funktion von ‚Multipart-Mails‘ realisiert. Die Bearbeiterin kann nun diesen Anhang öffnen und über das vorgefertigte Formular direkt bei Subito eine Bestellung tätigen. Dabei wird ebenso der Mechanismus des Cookies zur temporären Zwischenspeicherung benutzt wie bei der Datenbankanbindung.

Die Bestellung kann, abgesehen von der Anmeldung am Subito-System, ohne eine einzige zusätzliche Eingabe von seiten der Bearbeiterin auskommen. Das Aufheben des Medienbruches und die Verknüpfung der Systeme stellt einen wegweisenden Schritt in Richtung zukünftiger Systeme dar.

Das System speichert jede Bestellung lokal auf dem Web-Server im XML-Format ab. Diese gespeicherten Daten werden zur Zeit noch nicht weiterverarbeitet.

Ausblick

Für das Fernleihsystem im DKFZ wäre zu überlegen, ob die Anbindung von Internet-Literatur-Recherche-Systemen wie PubMed




- Neuer Bestellvorgang in der Zentralbibliothek des DKFZ:
- 1) Der Benutzer füllt das Web-Formular aus.
 - 2) Er bekommt nach dem Absenden eine Bestätigungsmail.
 - 3) Der Fernleihbestellauftrag wird in erweiterter Form an die Fernleihstelle gemailt.
 - 4) Die Bibliothek bestellt den Artikel bei Subito.
 - 5) Subito gibt den Auftrag weiter an eine Bibliothek im Subito-Lieferantenverbund. Diese Bibliothek im Subito-Lieferantenverbund hat der Besteller zuvor ausgewählt.
 - 6) Diese liefert den Artikel in gewünschter Form (Email, Fax oder Kopien) an die Fernleihstelle des DKFZ weiter.
 - 7) Die Fernleihstelle gibt den Artikel an den auftraggebenden Benutzer weiter und stellt gegebenenfalls die Kosten dem Institut in Rechnung.

Abb. 2 Neuer Fernleihvorgang im DKFZ

Interlibrary Loan Request	
Author	Schneider G, Senger B.
Titel	Coffee beans as a natural test food for the evaluation of th
Journal	J Oral Rehabil
Year	2001
Volume	28(4)
Pages	342-8
Name	Andreas Test
E-Mail	a.bohne@dkfz-heidelberg.de
Cost Center (Kostenstelle)	XYZ
Type of order	<input checked="" type="checkbox"/> Via Email (not always possible) <input type="checkbox"/> Normal copies <input type="checkbox"/> Fast order (please call additionally Frau Sitek via phone: 42-3661)
<input type="button" value="Order Now"/> <input type="button" value="New Article"/> <input type="button" value="Clear All"/>	

Abb. 4 Web-Formular zur online-Eingabe von Fernleihbestellungen

RechercheZeitschriften


Stichwort (Titel)

ISSN

CODEN

Sonstige:

Abb. 5 Web-Formular bei Subito. Der Cookie-Mechanismus setzt automatisch den Wert der eingegebenen Journal-Abkürzung ein und selektiert, dass im Bereich Zeitschriftentitel gesucht werden soll.

Dokumentbestellung bei der Deutsche Zentralbibliothek für Medizin

Signatur: Zs.A 991
 Zeitschriftentitel: Journal of oral rehabilitation
 Erscheinungsort: Oxford [u.a.]
 Verleger: Blackwell
 ISSN: 0305-182X
 CODEN: JORHB
 Dokumentennummer: 187079

Angaben zum Artikel:

Bitte geben Sie an, welchen Aufsatz Sie bestellen möchten und geben Sie das Erscheinungsjahr der Zeitschrift an:

Autor:	<input type="text" value="Schneider G, Senger B."/>	
Titel:	<input type="text" value="Coffee beans as a natural test food for"/>	
Jahrgang/Heft:	<input type="text" value="28(4)"/>	(Pflichtfeld)
Erscheinungsjahr:	<input type="text" value="2001"/>	(Pflichtfeld)
Seitenangaben:	<input type="text" value="342-8"/>	(Pflichtfeld)

realisiert werden sollte. Zwar gibt es bei PubMed Bestrebungen, Suchsystem und Bibliotheksbestände miteinander abzugleichen, aber dieses Projekt steckt noch in den Kinderschuhen und beschränkt sich erst einmal nur auf Nordamerika. Eine allgemeine Schnittstelle zum Übertragen von ausgewählten PubMed-Artikeln zu einem eigenen System existiert noch nicht.

Eine solche Schnittstelle kann aber sehr einfach über einen anderen Weg als den über die Webtechnologien realisiert werden. Man könnte einen Druckertreiber entwickeln, der, statt die Daten auf einen Drucker auszugeben oder in eine Datei umzuleiten, sie als Fernleihauftrag an die Bibliothek absendet. Unter Linux wird gerade ein solch rudimentärer Druckertreiber zum Testgebrauch entwickelt. Somit würde der Benutzer seine Recherche in PubMed tätigen und könnte dann die Artikel, die ihn interessieren, als Fernleihe (Bestellung) bei der Bibliothek aufgeben. In diesem Zusammenhang muss vorher genau überlegt werden, ob man einem Benutzer ein System an die Hand gibt, wo er mit einem Klick 10, 20, 50 oder mehr kostenpflichtige Fernleihen in Auftrag geben kann.

Danksagung

Ganz herzlichen Dank an Frau Dagmar Sitek. Frau Sitek koordiniert die Fernleihaufträge im DKFZ und war eine große Hilfe bei der Beantwortung von bibliotheksrelevanten Fragen.

Anschrift der Autoren:

Dr. Andreas Bohne
 Zentrale Spektroskopie
 Deutsches Krebsforschungszentrum
 Heidelberg
 Im Neuenheimer Feld 280
 D-69120 Heidelberg
 <<http://www.dkfz.de/spec/>>
 <a.bohne@dkfz.de>

Jörg Schwiemann
 Subito AG – Geschäftsstelle –
 Kurt-Schumacher-Damm 12
 D-13405 Berlin
 <<http://www.subito-doc.de>>
 <schwiemann@subito-doc.de>

Abb. 6 Web-Formular bei Subito. Der Cookie-Mechanismus setzt automatisch die Werte Autor, Titel, Heft, Jahr und Seiten ein