

Einsatz von BibTip als Recommendersystem im Bibliothekskatalog

Michael Mönnich, Marcus Spiering

Im Folgenden wird die Dienstleistung BibTip (<http://bibtip.org>) beschrieben, die einem Bibliothekskatalog eine Empfehlungskomponente hinzufügt. Dieser innovative Dienst ist ein wichtiger Baustein in der Entwicklung von bibliothekarischen Katalogen in Richtung des Web 2.0 und wird von der Universitätsbibliothek Karlsruhe als Dienst für andere Bibliotheken angeboten.

Recommendersysteme allgemein

Im Internet findet man zunehmend Systeme, die als Recommendersysteme bezeichnet werden. „Recommendation“ bedeutet Empfehlung. „Recommender“ ist derjenige, der die Empfehlungen ausspricht. Jeder, der schon einmal von einem Bekannten ein gutes Lokal oder einen guten Kinofilm empfohlen bekommen hat, nahm also schon einmal einen „Recommender“ in Anspruch.

Automatisch arbeitende Recommendersysteme erstellen Empfehlungen ohne die Einbeziehung von Menschen und werden insbesondere bei Onlineshops oft eingesetzt. Das unter Bibliothekaren bekannteste Beispiel dürften die Empfehlungen beim Online-Buchhändler Amazon (<http://www.amazon.de>) sein: Fast jede Titelaufnahme enthält dort einen Link „Kunden, die Bücher von ... gekauft haben, haben auch Bücher dieser Autoren gekauft: ...“. Diese Empfehlungen entstehen durch die Beobachtung und die Auswertung des Kaufverhaltens der Kunden bei Amazon.

Auch in Bibliotheken bieten sich vielfältige Einsatzmöglichkeiten für Recommender: Sie können nicht nur die Nutzer bei der Literaturrecherche unterstützen, sondern auch den Bibliothekaren beim Bestandsaufbau Informationen bieten. Auch die Sacherschließung kann durch einen Recommender sinnvoll ergänzt werden. Aus bibliothekarischer Sicht kann man Recommender als eine spezielle Form von Kataloganreicherung betrachten oder auch als einen Web-2.0-Baustein, da die Daten indirekt durch die Nutzer erzeugt werden und die Einbindung in den Katalog mittels eines Mashups (vergleichbar Google-Maps) erfolgt – zwei wesentliche Aspekte von Web-2.0-Anwendungen.

Das Karlsruher Recommendersystem BibTip

An der Universität Karlsruhe wurden im Zeitraum von 2002 bis 2007 mehrere DFG-Projekte durchgeführt, welche die Entwicklung von Recommendersystemen für den Einsatz in Bibliotheken zum Gegenstand hatten. Daraus ist BibTip hervor-

gegangen. Projektpartner waren dabei die Universitätsbibliothek Karlsruhe und das Institut für Informationswirtschaft und -management von Prof. Dr. Andreas Geyer-Schulz an der Fakultät für Wirtschaftswissenschaften (<http://www.em.uni-karlsruhe.de/research/projects/reckvk/>). Im Institut wurden die Algorithmen und die technischen Grundlagen von BibTip entwickelt. Die Bibliothek war für die Einbindung in den Katalog, die Erfassung des statistischen Datenmaterials und die Entwicklung der Dienstleistung BibTip verantwortlich. Das Projekt war so erfolgreich, dass es im Auftrag der DFG von den Projektnehmern im Dezember beim Fall 2007 Task Force Meeting der Coalition for Networked Information in Washington DC präsentiert wurde.

Bei BibTip handelt es sich um einen verhaltensbasierten Recommender. Dieser Typus von Recommenderdiensten basiert auf der – im Fall von BibTip anonymisierten – Beobachtung von Nutzerverhalten und der statistischen Auswertung dieser Daten. Im Internet-Handel ergeben sich die Nutzungsdaten aus Kaufvorgängen oder aus den Klicks auf Links in Webseiten. Im Falle von BibTip sind es die Aufrufe von Volltitelanzeigen im Online-Katalog.

Grundsätzlich arbeiten alle verhaltensbasierten Recommender so, dass aus der statistischen Auswertung der durch den Nutzer erzeugten Daten automatisch Empfehlungen erzeugt werden. Folgende Größen sind dabei relevant: Kunden, Produkte und Warenkörbe. Zur Generierung von Empfehlungen werden dabei die Produktpaare in Warenkörben von Kunden betrachtet. Die Fähigkeit, zufällig gemeinsam gekaufte Produktpaare sicher zu identifizieren, ist entscheidend für die Qualität des Recommenders. Zufällig gemeinsam gekaufte Produkte dürfen zu keiner gegenseitigen Empfehlungen führen, nicht zufällig zusammen gekaufte Produkte sollen empfohlen werden. Die Unterscheidung zwischen zufällig und absichtlich zusammen gekauften Produktpaaren trifft das Recommendersystem aufgrund von Statistiken. Entscheidend für die Qualität der erzeugten Empfehlungen sind die zugrunde liegenden statistischen Verfahren. Diese sind je nach Einsatzgebiet unterschiedlich, ihre Leistungsfähigkeit hängt davon ab, wie gut sie in der Lage sind, aus der großen Menge von Klicks die relevanten von den irrelevanten zu trennen.

Das algorithmische Verfahren, auf dem BibTip basiert, ist besonders robust gegenüber Störungen und speziell für die Daten in Bibliothekskatalogen entwickelt worden. Grundlage des Karlsruher Recommendersystems bildet die auf Andrew Ehrenberg zurückgehende Repeat-Buying-Theorie. Hierbei handelt es sich um eine beschreibende Theorie zum Konsumentenverhalten, die auf der Analyse von Warenkörben aufsetzt. Dabei wird der „Logarithmic Series Distribution“-Algorithmus verwendet, der eine 100%ige Automatisierung mit einer hohen Präzision verbindet. Literatur zu der zugrundeliegenden Theorie findet man unter <http://www.em.uni-arllsruhe.de/research/projects/reckvk.>

Alle verhaltensbasierten Recommenderdienste arbeiten umso besser, je größer die Datenbasis ist, denn nur bei einer ausreichend großen Anzahl von Transaktionen sind die statistischen Daten so aussagekräftig, dass brauchbare Empfehlungen entstehen. Also kann BibTip nur bei ausreichend stark frequentierten Systemen sinnvoll eingesetzt werden. Diese Voraussetzung sollte indes für die meisten Bibliothekskataloge – zumindest in Universitäts- und Hochschulbibliotheken – erfüllt sein.

BibTip bietet der Bibliothek einige Vorteile:

- Der Betrieb ist sehr kostengünstig, da es sich um ein rein maschinelles Verfahren handelt. Personal wird von Seite der Bibliothek nicht benötigt.
- Die Benutzer haben einen sichtbaren Mehrwert, der in Zeiten von Web 2.0 von vielen Benutzern auch erwartet wird.
- Die Empfehlungen sind stets völlig objektiv, da sie nur auf Statistiken beruhen.
- Empfehlungen veralten nie. Im Gegensatz zu einer Systematik werden sie laufend neu berechnet, das heißt dynamisch angepasst.
- Für den Bibliothekar ist interessant, dass die Anzahl der Empfehlungen zugleich ein Maß für die Nutzung eines Titels ist.
- Empfehlungen sind völlig medienneutral, das heißt, es können auch zu Videos, Audiodaten, Zeitschriftenaufsätzen etc. Empfehlungen ausgesprochen werden.

The screenshot shows the library catalog interface for Universität Karlsruhe (TH). At the top, there is a navigation bar with links for Home, Mein Konto, Kataloge, Digitale Bibliothek, Lieferdienste, Fachgebiete, and Infos. Below this is a search results section titled 'Suchergebnis'. The search query is 'Uni-Katalog' with a search term 'find nd=6914364'. The search results show a book by Murray Wrobel, 'Elsevier's dictionary of fungi and fungal plant diseases'. The book details include the language (Latin, English, German, French, and Italian), the compiler (M. Wrobel and G. Creber), the publisher (Elsevier, 1998), and the ISBN (0-444-82774-9). The bibliography and keywords are also listed. On the right side of the search result, there is a 'BibTip' link and a list of actions: Ausleihe, weitere Literatur, Mehr zu diesem Titel, Titelfprofil exakt, Zur Auswahl hinzufügen, Empfehlungen, Bewertung abgeben, Rezension schreiben, Meine Rezensionen, Als Favorit ablegen, and Meine Favoriten.

Abb. 1: Volltitelanzeige im Katalog der UB Karlsruhe mit BibTip-Link

Universität Karlsruhe (TH) | Universitätsbibliothek
Forschungsuniversität · gegründet 1825

Home | Mein Konto | Kataloge | Digitale Bibliothek | Liefersdienste | Fachgebiete | Informationen

Uni-Katalog: Empfehlungen zu Ihrem Suchergebnis

Dokument: nd=6914364
Katalog: UB Karlsruhe

[Neue Suche](#) [Suchergebnisse](#)

Ihr Suchergebnis

Eisevier's dictionary of fungi and fungal plant diseases / Wrobel, Murray; Creber, Geoffrey . 1998

Empfohlene Dokumente zu Ihrem Suchergebnis

Der links stehende Titel wurde auch zusammen mit folgenden Titeln aufgerufen:

(Anzahl der gemeinsamen Benutzungen in Klammern.)

1. Pilze / Laessøe, Thomas; Fletcher, Neil , 1999, (19) 📖 📄
2. The fungi / Carlile, Michael J.; Watkinson, Sarah C. , 1996, (17) 📖 📄
3. Schimmelpilze / Mücke, Wolfgang; Lemmen, Christa , 1999, (17) 📖 📄
4. Schimmelpilze / Reiß, Jürgen , 1998, (16) 📖 📄
5. Bacteria in biology, biotechnology, and medicine / Singleton, Paul , 1999, (16) 📖 📄
6. Aspergillus / Martinelli, Sylvia D. , 1994, (15) 📖 📄
7. Pilzgeschichten / Kothe, Hans; Kothe, Erika , 1996, (15) 📖 📄
8. Molecular fungal biology / Oliver, Richard P. , 1999, (15) 📖 📄




Abb. 2: BibTip-Empfehlungsliste

Einsatz von BibTip

Grundlage von BibTip ist also die statistische Auswertung der Katalognutzung. Ein anderer Ansatz wäre es, die Ausleihfrequenz von Büchern für Empfehlungen zu nutzen. Der BibTip-Ansatz bietet demgegenüber jedoch einige gewichtige Vorteile:

- Es fallen mehr Daten an, da meist mehr im Katalog recherchiert wird, als wirklich ausgeliehen.
- Auch Präsenzbestände und Literatur in Institutsbibliotheken werden erfasst.
- Die Verfügbarkeit hat keinen verfälschenden Einfluss auf die Empfehlungen.
- Es treten keine Probleme mit Datenschutz auf, da nur anonyme OPAC-Sessions gespeichert werden. Würde man Ausleihdaten nutzen, so müssten diese zusätzlich anonymisiert werden.

Bei BibTip wird also der oben beschriebene Recommenderansatz so umgesetzt, dass dem Kunden der OPAC-Benutzer und dem Produkt die Titelaufnahme entspricht. Als Produktnummer dient die OPAC-Identnummer oder ersatzweise die ISBN, der Warenkorb ist durch die Menge der während einer OPAC-Sitzung betrach-

teten Titelaufnahmen gegeben. Voraussetzung für BibTip ist also ein leistungsfähiger und flexibler OPAC, in den sich das Recommendersystem nahtlos integrieren lässt. BibTip erzeugt Empfehlungen nur für diejenigen Titel, die im Katalog wirklich angeklickt werden. Der Aufruf von Katalogdaten im Verhältnis zum Gesamtbestand folgt in der Regel einer Pareto-Verteilung, das heißt ca. 80% der Aufrufe entfallen auf ca. 20% der vorhandenen Titel. In Karlsruhe, wo BibTip seit 2002 in Betrieb ist, liegen für den Uni-Katalog (<http://www.ubka.uni-karlsruhe.de/katalog>) inzwischen 200.000 Empfehlungslisten vor. Da der Katalog insgesamt 950.000 Titel enthält, ist somit eine Abdeckung von über 80% erreicht. Die Sortierung innerhalb der Empfehlungsliste erfolgt nach der Anzahl der gemeinsamen Benutzungen, das heißt, die meistgefragten Titel stehen oben.

Bevor die ersten Empfehlungen angezeigt werden, muss, wie beschrieben, zunächst eine reine Beobachtungsphase durchlaufen werden, in der nur Daten gesammelt und analysiert werden. Erst wenn eine ausreichend große Zahl von Empfehlungen vorliegt, ist es sinnvoll, diese anzuzeigen. Die Dauer der Beobachtungsphase ist um so kürzer, je mehr Daten anfallen, das heißt je intensiver der Katalog benutzt wird. Daneben haben natürlich auch die Anzahl der pro Sitzung betrachteten Titel und deren Diversizität Einfluss auf die Länge der Beobachtungsphase. Eine genaue Prognose ist also sehr schwierig, man kann aber davon ausgehen, dass es meist mehrere Monate dauern wird, bis genügend statistisches Material für valide Empfehlungen vorliegt. Den Aufbau kann man beschleunigen, indem man die Statistikdaten mehrerer Bibliotheken „poolt“. Dies ist vor allem dann möglich, wenn es sich um Bibliotheken mit einem sehr ähnlichen Titelbestand und einer vergleichbaren Klientel (und vergleichbarem Nutzerverhalten) handelt, zum Beispiel bei Stadtbibliotheken.

BibTip als Dienstleistung für die eigene Bibliothek nutzen

Wie oben erläutert erweitert BibTip die Funktionalität von Online-Bibliothekskatalogen um Empfehlungen, die durch das anonymisierte Beobachten des Benutzerverhaltens bei der Katalog-Recherche und dessen statistischer Analyse erzeugt werden. Von BibTip erzeugte Empfehlungen erscheinen als Hyperlinks in der Volltitelanzeige des OPACs und führen den Benutzer zu inhaltlich verwandten Titeln.

Die Universitätsbibliothek Karlsruhe bietet BibTip interessierten Bibliotheken seit Sommer 2007 als eine Dienstleistung an. Sie ist so konzipiert, dass von Seiten des Kataloganbieters keine Installation zusätzlicher Software notwendig ist, sozusagen als eine Art von Webservice. Die statistische Datenauswertung, die Erzeugung und die Verwaltung der Empfehlungen findet vollständig auf Servern in Karlsruhe statt. Zur Nutzung von BibTip ist – neben einem zuverlässigen Internetanschluss – lediglich eine Erweiterung der Volltiteldarstellung des Online-Biblio-

theyskataloges erforderlich. Im günstigsten Fall muss eine Zeile JavaScript-Code in den HTML-Code der Volltiteldarstellung eingefügt werden, um die beschriebene Funktionalität zur Verfügung zu stellen. Die Funktionsweise der Einbindung von BibTip ist unter <http://www.bibtip.org/integration.html> näher beschrieben.

Für die Nutzung erhebt die UB Karlsruhe ein festes jährliches Entgelt, das dazu verwendet wird, den laufenden Betrieb von BibTip zu finanzieren und die Software weiterzuentwickeln. Näheres hierzu regelt ein Dienstvertrag, den die Bibliothek mit der UB Karlsruhe abschließt. Bisher haben sich die Bibliotheken der Universitäten Freiburg, Braunschweig, TU Berlin, Stuttgart und Stuttgart-Hohenheim, die Badische Landesbibliothek, die Deutsche Nationalbibliothek sowie mehrere Hochschulbibliotheken in Baden-Württemberg für den Einsatz von BibTip entschieden. Sie werden nach und nach in die Beobachtungsphase eintreten; mit mehreren anderen Bibliotheken werden derzeit Gespräche geführt. Bei Interesse an BibTip wenden Sie sich bitte an die Autoren oder einfach per E-Mail an bibtip@ubka.uni-karlsruhe.de.