

bibnet.org – kooperative Referenzdatenbank für das Gesundheitswesen

bibnet.org – collaborative health care reference database

Abstract

Members of a Swiss, Austrian and German network of health care libraries built a collaborative and freely accessible article reference database: <http://bibnet.org/>.

The database contains primarily references of German health care journals. Since different libraries were cataloging articles on their own, the goal was to merge existing collections. As of November 2010 the database *bibnet.org* contains 45,000 article references from 400 journals. The records origin in the first place from the Rudolfinerhaus in Vienna and the Swiss Pro Senectute library, dating back till 1979. Additional records are being catalogued by participating libraries. Each participating library is responsible for an assigned set of journals.

The project was built using exclusively Open Source components. Vufind as a discovery layer, KOHA as cataloguing system for those libraries without a capable ILS, Drupal as content management system and Doctor-Doc as linkresolver, ILL system and through its DAIA interface as knowledge database to indicate print holdings availability for the participating libraries.

bibnet.org shows the potential libraries can have if they start to cooperate.

Keywords: reference database, libraries network, Open Source, MARC21

Zusammenfassung

Eine Arbeitsgruppe des Vereins „Netzwerk Fachbibliotheken Gesundheit“ (CH) hat unter der Adresse <http://bibnet.org/> zusammen mit dem Rudolfinerhaus in Wien (A) eine kooperative und frei zugängliche Referenzdatenbank für das Gesundheitswesen geschaffen.

Nachgewiesen werden Referenzen primär aus dem pflegerischen Bereich stammender und überwiegend deutschsprachiger Zeitschriftenartikel. *bibnet.org* führt auf einer zentralen Plattform vorhandene Katalogisate verschiedener Bibliotheken zusammen.

Aktuell enthält die Datenbank rund 45.000 Datensätze aus über 400 ausgewerteten Zeitschriften. Diese stammen aus den Beständen des Rudolfinerhauses in Wien und der Pro Senectute Bibliothek Schweiz und gehen bis ins Jahr 1979 zurück. Weitere Datensätze werden durch Beteiligung zahlreicher Bibliotheken an der fortlaufenden Katalogisierung hinzugefügt. Jede teilnehmende Bibliothek übernimmt dabei die Verantwortung für die Auswertung der ihr zugeteilten Zeitschriften.

Sämtliche verwendeten technischen Systeme basieren auf Open Source-Lösungen: Als Suchsystem kommt Vufind zum Einsatz. Als Katalogisierungssystem für Bibliotheken, die über keine MARC-kompatiblen Systeme verfügen, steht eine Instanz von KOHA zur Verfügung. Als Content-Management-System für den Internetauftritt wird Drupal eingesetzt. Als offenes Linkresolver- und Bestellsystem dient Doctor-Doc und über eine DAIA-Schnittstelle wird die Verfügbarkeit der Artikel in den teilnehmenden Bibliotheken angezeigt.

Das Projekt *bibnet.org* hat als kooperatives Werk und frei zugängliche Datenbank Modellcharakter für das Bibliothekswesen. Es wird sichtbar,

Markus Fischer¹
Stefan Kandra²
Veronika Kleibel³
Maike Krone⁴
Susanne Mayer^{5,6}
Erika Niedermann⁷
Dieter Sulzer²

1 Solothurner Spitäler AG,
Solothurn, Schweiz

2 Pro Senectute, Bibliothek
und Dokumentation, Zürich,
Schweiz

3 Rudolfinerhaus Wien,
Österreich

4 Bildungszentrum Gesundheit
Basel-Stadt, Münchenstein,
Schweiz

5 Careum Stiftung, Zürich,
Schweiz

6 FH Südschweiz,
Physiotherapie Landquart,
Schweiz

7 Höhere Fachschule
Gesundheit Luzern, Schweiz

was in der Vernetzung bestehender bibliothekarischer Ressourcen möglich wird.

Schlüsselwörter: Referenzdatenbank, bibliothekarische Vernetzung, Open Source, MARC21

Entstehung von bibnet.org

Im Jahr 2005 gründeten die Pflegebibliotheken der Schweiz das „Netzwerk Pflegebibliotheken“, um den Informationsaustausch untereinander zu verbessern. 2009 schloss es sich mit dem Netzwerk der Spitalbibliotheken „SpiBiNet“ zum Verein „Netzwerk Fachbibliotheken Gesundheit“ zusammen.

Im Auftrag des Vereins kümmert sich seit 2009 eine Arbeitsgruppe um die Schaffung einer kooperativen Referenzdatenbank für (deutschsprachige) Zeitschriften des Gesundheitsbereiches.

Ausgangslage

Die Arbeitsgruppe stellte fest, dass zahlreiche Bibliotheken, unabhängig voneinander, Zeitschriften auswerten und Referenzdatensätze für ihre lokalen Kataloge erstellen.

Diese Daten sind jeweils nur einem relativ begrenzten Kreis an Personen zugänglich. Der Impact der geleisteten Arbeit ist gering.

Zudem existiert keine systematische Auswertung von deutschsprachigen (Pflege-)Zeitschriften. Die bestehenden Suchmöglichkeiten nach Publikationen in dieser Literatur sind relativ begrenzt: CCMed, Gerolit und das kostenpflichtige Angebot Carelit gehören zu den wenigen Suchmöglichkeiten. Erschwerend kommt hinzu, dass die Verlagsangebote in diesem Literaturbereich technisch sehr oft nicht den Anforderungen der Zeit entsprechen: zahlreiche Zeitschriften sind nicht online, oft sind nur Inhaltsverzeichnisse verfügbar, und falls Inhalte online sind, werden ganze Heftausgaben als PDF eingestellt anstelle der üblichen PDFs auf Artekelebene. Unter diesen Voraussetzungen sind die bestehenden Angebote auf Artekelebene in der Regel nicht Open URL-fähig auf, sondern bestenfalls auf Zeitschriftenebene. Identifikatoren wie DOIs sind kaum vorhanden, teilweise fehlen sogar ISSN-Nummern.

Zielsetzungen

Die Arbeitsgruppe hat sich zum Ziel gesetzt, die bestehenden lokalen Referenzdaten zusammenzuführen und in einem Suchsystem zur Verfügung zu stellen, das den Suchgewohnheiten von Usern gerecht wird. Folgende Anforderungen wurden aufgestellt:

- Frei zugänglich
- Betrieb und Verwaltung durch und für Bibliotheken
- Einsatz von Open Source-Lösungen
- MARC21 als Datenformat
- AACR2 als Katalogisierungsregelwerk

- Das System soll Open URL-fähig sein
- Anzeige der Verfügbarkeit auf Artekelebene

Eingesetzte Software

Da der Arbeitsgruppe nur beschränkte Geldmittel zur Verfügung stehen, mussten alle eingesetzten Softwarelösungen auf Open Source basieren. *bibnet.org* besteht aus 5 Softwarekomponenten:

- Drupal: Internetauftritt (CMS)
- KOHA: Katalogisierung
- *bibnet.org* – dedupe utilities: maschinelle Datenkontrolle und Deduplikation
- Vufind: Suchoberfläche
- Doctor-Doc: ISSN und Bestandesregister für die Verfügbarkeitsanzeige (DAIA) und als Linkresolver unter Verwendung der Dienstleistungen der EZB/ZDB (JOP)

Drupal

<http://drupal.org/>

Als Internetpräsenz kommt Drupal zum Einsatz. Drupal dient der Koordination von Bibliotheken und Zeitschriften.

KOHA

<http://koha-community.org/>

Koha ist ein weit verbreitetes Open Source-Bibliothekssystem (ILS).

Für teilnehmende Bibliotheken ohne MARC-kompatible Bibliothekssysteme wird online eine Installation von KOHA zur Verfügung gestellt. Für die teilnehmenden Bibliotheken sind keine lokalen Installationen nötig.

In Koha wurde für jede Zeitschrift eine Katalogisierungsvorlage erstellt, um einfaches Katalogisieren zu ermöglichen: sich wiederholende Angaben wie Zeitschriftentitel und ISSN sind jeweils bereits vordefiniert.

Vufind

<http://vufind.org/>

Vufind ist eine auf bibliothekarische Bedürfnisse zugeschnittene Suchoberfläche (Discovery layer).

Als Recherchetool lässt diese Lösung kaum Wünsche offen. Es stehen u.a. folgende Funktionen zur Verfügung:

- Facetted Browsing (geführte Suche)
- Fuzzy-, Proximity-, trunkierte Suche
- Zitier- und Exportfunktionen
- Über Tags und Kommentare können Artikel angereichert werden (User Created Content)

- Erstellen personalisierter und öffentlicher Literaturlisten
- Bibliographische Angaben per E-Mail verschicken
- Prüfung von Online-Beständen via Open URL

Von uns hinzugefügte Funktionen umfassen u.a.:

- Trefferübersicht
 - automatische Verfügbarkeitsanzeige für eine spezifische Institution (IP-basiert)
- Detailansicht
 - Anzeige aller Institutionen in denen der Artikel vorhanden ist
- Location Switcher
 - über ein Drop-Down-Menü kann die Verfügbarkeitsansicht einer spezifischen Institution auch ausserhalb eines registrierten IP-Bereiches ausgewählt werden.

Technisches

Um die Anzeige der Verfügbarkeit zu ermöglichen, ist es notwendig über maschinenlesbare Artikelangaben oder Identifikatoren zu verfügen.

Bei Artikelreferenzen reduziert sich das in der Regel auf:

- ISSN
- Jahrgang
- Heftnummer
- Anfangsseite/Schlussseite

Eindeutige Identifikatoren wie DOIs, PMIDs, ZDB-IDs und andere Identifikatoren sind im Bereich der ausgewerteten Literatur nur teilweise vorhanden und nicht unmittelbar mit den Beständen verknüpfbar.

MARC21

Sehr erstaunlich ist, dass MARC21 keine separaten Felder für Artikelreferenzen vorsieht. Diese Angaben sind zwingend, um grundlegende Funktionen wie Open URL zu realisieren. MARC21 speichert diese Angaben im Freitext-Feld 773g:

773 \$g 50(2010), no. 3, p. 352-362

Freitext-Felder müssen in die Einzelelemente zerlegt werden, bevor die Angaben maschinenlesbar verwertet werden können. Ausserdem garantieren sie keine einheitliche Datenstruktur, was für die zuverlässige Auswertung von Referenzangaben unabdingbar wäre. Die Ausgangslage wurde zusätzlich erschwert, da die Daten aus unterschiedlichen Quellen mit unterschiedlichen „Regeln“ für die Gestaltung von 773g stammen. Es wurde schnell klar, dass diese Angaben nicht direkt in einem Produktionssystem geparkt und verwertet werden können.

Leider gibt es keine offizielle Lösung, um dieses mangelhafte Design der Datenstruktur des MARC-Formats anzugehen. Es ist anzunehmen, dass weltweit hunderttausende bibliographischer Artikeldaten im MARC-Format vorliegen, die ausschliesslich lokal oder in Verbänden, also nur in ein und demselben ILS, zugänglich gemacht werden

können. Diese Daten sind üblicherweise direkt mit Exemplardaten verknüpft. Die Kombination von ungenügender Datenstruktur und der Verknüpfung mit Exemplardaten erschwert das Zusammenführen von Datenbeständen in einem externen Kontext.

Bei der Suche nach einer Lösung für dieses Problem sind wir auf ein Discussion Paper von 2003 der Library of Congress (<http://www.loc.gov/marc/marbi/2003/2003-dp01.html>) gestoßen und haben die Empfehlung 4.1 umgesetzt:

Dazu wurden drei zusätzliche Unterfelder für den Jahrgang (773v), die Heftnummer (773l) und für die Anfangsseite (773q) definiert.

Danach konnte das summarische Feld in die neuen Unterfelder geparkt und die ergänzten Daten in Vufind importiert werden.

Da sämtliche Bibliothekssoftware rund um Formate wie MARC21 aufgebaut ist, ist die Nutzung der Unterfelder 773v/773l/773q standardmässig nicht vorgesehen. In der Folge mussten wir Vufind entsprechend anpassen:

- Darstellung der Trefferübersicht/Detailansicht/„Meine Favoriten“
- Ergänzung Open URL
- Zitierfunktion
- Export nach Refworks/Endnote

Da Vufind sehr sauber programmiert und klar strukturiert ist, waren diese Änderungen mit relativ kleinem Aufwand realisierbar.

Verfügbarkeitsanzeige – DAIA

Für die Anzeige der Verfügbarkeit konnten wir auf einen existierenden DAIA-Treiber der TU Hamburg-Harburg zurückgreifen, den wir noch mit IP-basierten Funktionen erweitert haben.

DAIA (Document Availability Information API) ist ein Protokoll, um die Verfügbarkeit von Dokumenten abzufragen. DAIA wurde in Zusammenarbeit von der Verbundzentrale des GBV, der HeBIS-Verbundzentrale und dem Beluga-Projekt der Hamburger Bibliotheken entwickelt: http://www.gbv.de/wikis/cls/Verf%C3%BCgbarkeitsrecherche_mit_DAIA

DAIA funktioniert unter der Annahme, dass ein eindeutiger Identifikator (z.B. Systemnummer) für die Abfrage vorhanden ist.

Artikel können in unserem Umfeld, wie aufgezeigt, kaum anhand eines eindeutigen Identifikators bestimmt werden. Ausserdem gibt es, anders als bei Verbundsystemen, keine typische Beziehung über eine ID zwischen einem Datensatz und einem Bestandeseintrag.

Deshalb werden die Bestandsanfragen über Open URL unter Verwendung der Referenzangaben an unseren DAIA-Server (<http://www.doctor-doc.com/>) geschickt. Optional kann eine IP mitgeschickt werden, um die Ansicht einer spezifischen Institution zu erfragen.

Doctor-Doc

<http://sourceforge.net/projects/doctor-doc/>

Als Bestandesregister dient Doctor-Doc. Bestände können über einfache Listen importiert und exportiert werden.

Um eine Bestandsanfrage beantworten zu können, sind für Artikel grundsätzlich nur minimale Angaben erforderlich:

ISSN und Start-/Endjahr reichen um einen Bestand abzubilden.

Doctor-Doc übernimmt zusätzlich ein internes ISSN-Mapping. Eine Zeitschrift kann durchaus über mehrere ISSN-Nummern verfügen (P-ISSN, E-ISSN, L-ISSN, CD-ROM). Doctor-Doc verfügt über eine Knowledge Database aller ISSN-Nummern und beantwortet eine Anfrage mit einer ISSN mit allen Bestandeseinträgen inkl. der „verwandten“ ISSN-Nummern unter Berücksichtigung des betreffenden Publikationsdatums. Verfügbarkeitsanfragen werden heftgenau beantwortet.

Dublettenkontrolle

<http://sourceforge.net/projects/bibnet/>

Da die Datenbestände aus unterschiedlichen Quellen stammen (KOHA, Rudolfinerhaus Österreich, Pro Senectute Schweiz), muss eine Dublettenkontrolle gewährleistet sein.

Sowohl KOHA als auch Vufind (über den integrierten SOLR-Index) verfügen über Deduplikationsverfahren. Es hat sich aber gezeigt, dass jeweils zu viel oder zu wenig dedupliziert wird.

Wir haben deshalb eine eigene Applikation „bibnet.org – dedupe utilities“ entwickelt, die heute als Open Source zur Verfügung steht.

Zum Einsatz kommt ein abgestuftes Deduplizierverfahren anhand der Referenzangaben kombiniert mit einer Näherungssuche im Titel anhand der „Damerau-Levenshtein Distance“ nach Entfernung der Leer- und Sonderzeichen im Titel (String-Normalisierung). Die „Damerau-Levenshtein Distance“ berechnet die Anzahl Unterschiede von zwei Vergleichstexten.

Ausblick und Herausforderungen

Die Arbeitsgruppe sieht folgende Entwicklungsschwerpunkte für die Zukunft:

- Qualitätskontrolle Katalogisate
- Automatisierung Importprozesse
- Bestellmöglichkeit von Artikelkopien (z.B. über Subito)
- Erstellen eines gemeinsamen Thesaurus/Verschlagwortung
- Erweiterung mit nicht deutschsprachigen Bibliotheken (Suisse Romande)
- E-Mail Alerting System
- Schnittstellen

Schlussfolgerungen

Das Projekt *bibnet.org* hat als kooperatives Werk und frei zugängliche Datenbank Modellcharakter für das Bibliothekswesen. Es wird sichtbar, was in der Vernetzung bestehender bibliothekarischer Ressourcen möglich wird.

Anmerkung

Interessenkonflikte

Die Autoren erklären, dass sie keine Interessenkonflikte in Zusammenhang mit diesem Artikel haben.

Korrespondenzadresse:

Markus Fischer

Solothurner Spitäler AG, Schöngrünstrasse 42, 4500 Solothurn, Schweiz

Bitte zitieren als

Fischer M, Kandra S, Kleibel V, Krone M, Mayer S, Niedermann E, Sulzer D. *bibnet.org – kooperative Referenzdatenbank für das Gesundheitswesen*. *GMS Med Bibl Inf*. 2010;10(3):Doc27.

DOI: 10.3205/mbi000210, URN: urn:nbn:de:0183-mpi0002103

Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/mbi/2010-10/mbi000210.shtml>

Veröffentlicht: 21.12.2010

Copyright

©2010 Fischer et al. Dieser Artikel ist ein Open Access-Artikel und steht unter den Creative Commons Lizenzbedingungen

(<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.de>). Er darf vervielfältigt, verbreitet und öffentlich zugänglich gemacht werden, vorausgesetzt dass Autor und Quelle genannt werden.