

Katalog.plus – OPAC 2.13

Teil 2: Datenbank-Empfehlungsdienst

[Friedrich Summann](#) & [Christian Pietsch](#)
[Universitätsbibliothek Bielefeld](#), [LibTec](#)
[Kolloquium Wissensinfrastruktur](#) am 14.12.2012



Ausgangslage

Fachdatenbanken für die Literaturrecherche sind umständlich zu finden:

1. www.ub.uni-bielefeld.de aufrufen →
2. *Fachinformation* klicken →
3. Fach aus Liste wählen →
4. *Fachdatenbanken (DBIS)* klicken →
5. evt. ein Fachgebiet auswählen.

Zu viele Klicks und zu weitab gelegen.

Nachtrag: Es gibt auch einen Link von der Hauptseite zu ub.unibi.de/dbis, aber ich hatte ihn übersehen. Das Auswahlproblem bleibt bestehen, ebenso die bisher fehlende Sichtbarkeit der Datenbanken innerhalb des OPAC.

Lösungsansatz

- Relevante Fachdatenbanken(-Listen) im [OPAC](#) empfehlen abhängig von der Suchanfrage!
→ [Empfehlungsdienst](#) (recommender system)
- Vorreiter beim Einsatz von Empfehlungsdiensten im OPAC: [SuUB Bremen](#)
- Suchanfrage → 1–3 Fachdatenbanken-Listen
- Direkte Lösung (NLP/KI) schwierig, langsam, unzuverlässig:
 - automatische Sprachverstehen (NLU) ist unausgereift
 - autom. Kategorisierung (ML) benötigt Trainingsdaten
 - Ontology-Matching ist aufwändig und nicht verlustfrei
 - verwendete Algorithmen sind zeit- und speicherintensiv

Lösung

- Indirekte Lösung (IR/BI) einfach, schnell, zuverlässig:
Das im Katalog gesammelte Wissen der Katalogisierer nutzen, indem die Metadaten der ersten (z.Z. 20) Suchergebnisse ausgewertet werden.
- Beispiel: Suchanfrage = »beauvoir«
ergibt Funde aus den Fachbibliotheken 16; 15; 05; 12; 15; 14; 16; 05; 16; 15; 15; 15; 16; 15; 15; 15; 15.
- naive Auswertung: 9 × FB 15; 4 × FB 16; 2 × FB 05; Rest 1 ×
- liefert ungenaue Ergebnisse (s. folgende Tabelle), weil einer Fachbibliothek mehrere Fachgebiete zugeordnet sein können

FB	DBIS-Fachgebiete	Erklärung
03	16	Wirtschaftswissenschaften
05	FF	Frauen- und Geschlechterforschung
07	20	Sportwissenschaft
10	30 02	Informatik Mathematik
11	15	Rechtswissenschaften
12	18 17 29	Soziologie, Politologie, Ethnologie
13	16	Wirtschaftswissenschaften
14	27 26 09	Archäol., Geschichte, Klass. Philologie, Kunst, Archit.
15	13 12 11 09 10 51	Allg. u. Vergl. Sprach- u. Lit.wiss., Angl., German. u. Skandinav., Klass. Philologie, Romanistik, Slawistik
16	19 21 22 23 25 53	Philosophie Psychol. Theol. Päd.+Medienw. Musik
17	01	Physik
18	03	Chemie
19	05 48 52	Biologie
22	06 7	Geographie Geowissenschaften
23	GE	Gesundheitswissenschaften

Implementierte Auswertung

Die Auswertung der ersten Fundstellen nach Fachbibliotheken ist zu grob. Daher betrachten wir statt dessen die ersten beiden Stellen der Systemstelle/Signatur. Laut Systematik sind 400 dieser Präfixe an der UB Bielefeld in Verwendung.

- Systemstellen-Anfänge für das Beispiel »beauvoir«:
BJ VG AR IA VG DY BG AR AT VA VG VG BJ VG VG VG VG
- Ergebnis: Empfehlungen für je 5 Top-Datenbanken aus ...
 1. Fachgebiet: Romanistik (8 VG)
 2. Fachgebiet: Philosophie (1 BG + 2 BJ)
 3. Sammlung: Frauen- und Geschlechterforschung (2 AR)

Systematik-Exzerpt in Perl

```
my %systemstelle_fachgebiet_map = (
  'AA' => '54', # FB 01 Informationswissenschaft
  'AB' => '54', # FB 01 Informationswissenschaft
  'AC' => '54', # FB 01 Informationswissenschaft
  'AD' => '54', # FB 01 Informationswissenschaft
  'AF' => '_AF_', # FB 02 ZIF Allgemein / Fachübergreifend (Sonderfall)
  'AR' => 'FF', # FB 05 Frauen- und Geschlechterforschung <--- 2 x
# ...
  'BA' => '21', # FB 16.1 Philosophie
  'BB' => '21', # FB 16.1 Philosophie
  'BC' => '21', # FB 16.1 Philosophie
  'BD' => '21', # FB 16.1 Philosophie
  'BE' => '21', # FB 16.1 Philosophie
  'BF' => '21', # FB 16.1 Philosophie
  'BG' => '21', # FB 16.1 Philosophie <--- 1 x
  'BH' => '21', # FB 16.1 Philosophie
  'BI' => '21', # FB 16.1 Philosophie
  'BJ' => '21', # FB 16.1 Philosophie <--- 2 x
# ...
  'VA' => '10', # FB 15 - Romanistik
  'VB' => '10', # FB 15 - Romanistik
  'VC' => '10', # FB 15 - Romanistik
  'VD' => '10', # FB 15 - Romanistik
  'VE' => '10', # FB 15 - Romanistik
  'VF' => '10', # FB 15 - Romanistik
  'VG' => '10', # FB 15 - Romanistik <--- 8 x
# ...
);
```

Ausblick

Bedarf und Ressourcen vorausgesetzt, sind folgende Erweiterungen denkbar:

- Empfehlungen für bestimmte Datenbanken aus den 9867 vorhandenen, statt nur auf Listen von Top-Datenbanken zu verweisen
- Nutzung von in den Metadaten enthaltenen geographischen Angaben (z.B. XT für Römisches Reich, XA-DE-NW für Nordrhein-Westfalen)
- Weitere Vorschläge?

Bitte an die FachreferentInnen

Anzahl der Top-Datenbanken bei DBIS auf maximal 5 beschränken.

Alternative: Sortierung nach Nützlichkeit (absteigend).

Vielen Dank!

Präsentation gebaut mit [html5slides](#), einer freien Vorlage für Präsentationen in HTML5 von Google.