



Aufbau eines zentralen Indexes im GBV

Till Kinstler, kinstler@gbv.de

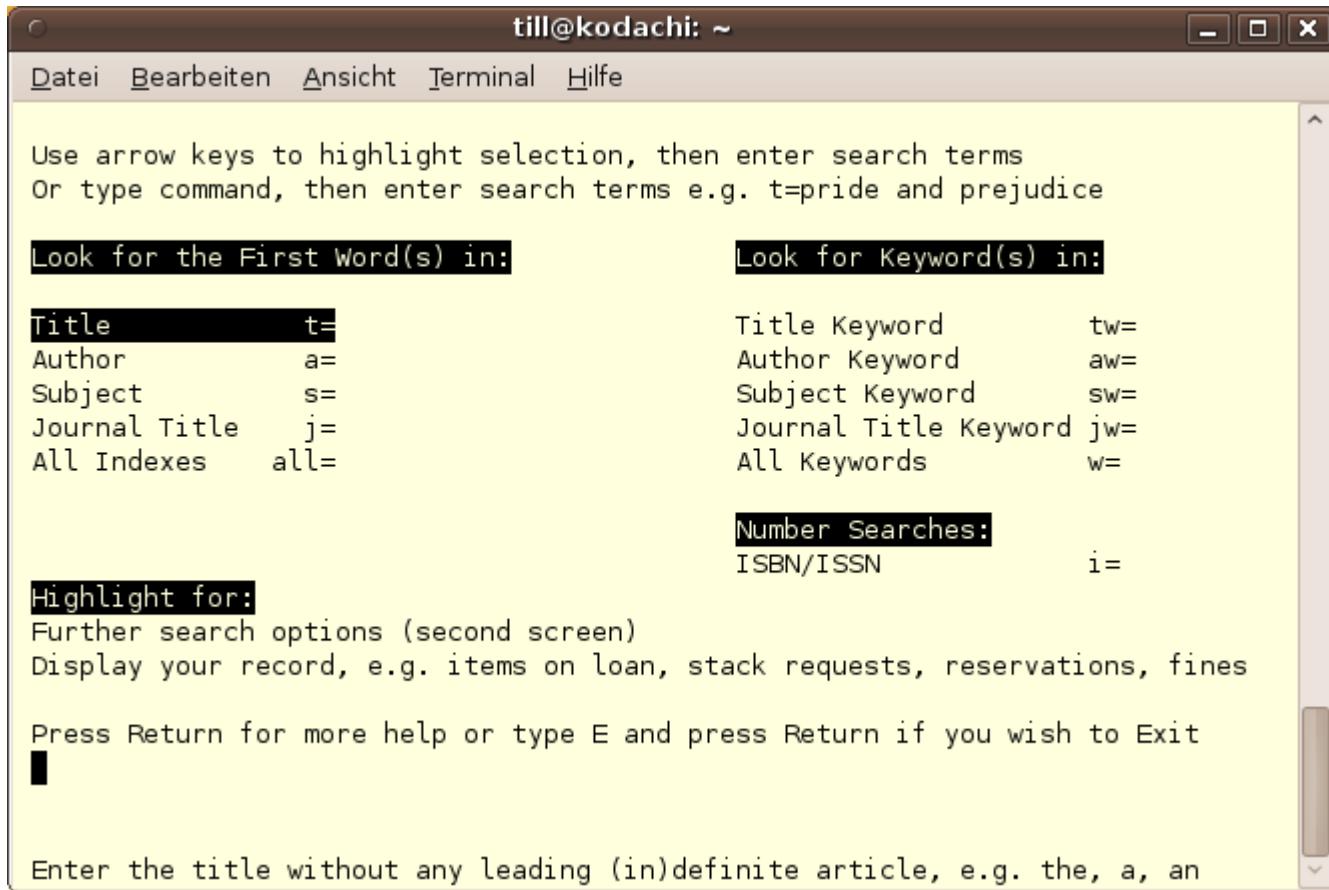
<http://twitter.com/tillk>

Verbundzentrale des GBV (VZG)

Wir haben im zentralen Metadatenmanagementsystem des GBVs (CBS) rund 100 Millionen interessante Datensätze: klassische OPAC-Daten (GVK), Online Contents (OLC), Nationallizenzen, ...

Und was machen wir damit?





search filter

search

Fill in one or more words in the search form below and add the desired settings.

[clear form](#)

all words [ALL]

and ▾

all subjects [ASB]

and ▾

author [AUT]

and ▾

title: keywords [TTL]

sort by year of publication ▾

year of publication

for example: 1948-1980 or 1948- or 1955

language

-- All languages -- ▾

country

-- All countries -- ▾

library

-- All libraries -- ▾

approximate search



material selection

[all](#)[none](#) Books Periodicals (non-online) Online periodicals Online resources (without periodicals) Articles Letters Audio visual Pictures Cartography Handwriting Music Sound Software Microfilm Miscellaneous

	Home >> Databases
Databases	<ul style="list-style-type: none">▶ GVK - GBV Union Catalogue <small>INFO</small>▶ OLC - Online Contents <small>INFO</small>▶ GVK-PLUS - GBV Union Catalogue with Online Contents <small>INFO</small>⊕ Online Resources - DOAJ, OLR, ZVDD <small>INFO</small>⊕ Online Contents Special Subject Collections (OLC-SSG) <small>INFO</small><ul style="list-style-type: none">⊕ Humanities⊕ History⊖ Engineering and Mathematics<ul style="list-style-type: none">▶ Online Contents - SSG Architecture and Civil Engineering <small>INFO</small>▶ Online Contents - SSG Environment <small>INFO</small>▶ Online Contents - SSG History of Technology <small>INFO</small>▶ Online Contents - SSG Mathematics and Informatics <small>INFO</small>▶ Online Contents - SSG Technology <small>INFO</small>⊕ Natural Sciences and Medicine⊕ Law, Economics and Social Studies⊕ Linguistics and Area Studies⊕ Special Subject Catalogues (Fach-OPAC)⊕ Nationally Licensed Databases⊕ Other Union and National Databases <small>INFO</small>⊕ Regional Catalogues⊕ Library Specific Subject Catalogues⊕ Library Catalogues - OPACs⊕ Databases for Public Libraries <small>INFO</small>⊕ Databases of K.G. Saur, Munich - AIDA, AKL-IKD, DBE, IBR, IBZ, IJBF, IJBK <small>INFO</small>⊕ Miscellaneous Databases⊕ Alphabetic list of all databases
Information	
GBV libraries	
ViFa & OLC-SSG	
Links	
Sitesearch	
Newsletter	
Mailinglists	
GBV-Wiki	

MetaSuche

Durchsuchen Sie mehrere Datenbanken gleichzeitig.

1. Bitte wählen Sie aus:

Schnellsuche-Sets

Kataloge

Nachschlagewerke
eJournal Pakete
Philosophie/Theol.
Geschichte
Physik/Mathematik
Bildung/Psychol.
Literatur/Sprache
Sozialwissenschaften
Recht/Wirtschaft
Chemie/Life Sciences
Geowissenschaften
Kunstwissenschaften
Gesundheitswissenschaften

Einfach Erweitert

3.

Alle Felder ▼ [] UND []

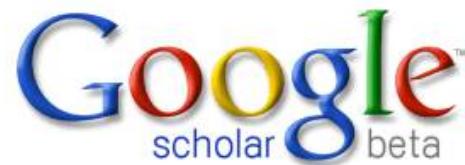
Alle Felder ▼ []

GO

2. Set: Kataloge enthält 12 E-Ressourcen

Alle auswählen/Auswahl aufheben

Name	Typ	Aktionen
Online-Katalog der UB Wien (OPAC)	Katalog	(i) +
Online-Katalog des Österreichischen Bibliothekenverbundes (OPAC)	Katalog	(i) +
GVK - Gemeinsamer Verbundkatalog	Katalog	(i) +
Library of Congress Online Catalog	Katalog	(i) +
COPAC - Verbundkatalog der Universitätsbibliotheken Großbritanniens und Irlands	Katalog	(i) +
NEBIS - Netzwerk von Bibliotheken und Informationsstellen in der Schweiz	Katalog	(i) +
SBN - Servizio Bibliotecario Nazionale	Katalog	(i) +
SUDOC - Système Universitaire de Documentation	Katalog	(i) +
Eurobuch.com	Suchmaschine	(i) +

 [Advanced Scholar Search](#)

Articles (include patents) Legal opinions and journals

Stand on the shoulders of giants

[Get email alerts for new publications](#)

[Go to Google Home](#) - [About Google](#) - [About Google Scholar](#)

©2010 Google

„In addition, we have also found that the **poor usability, high complexity, and lack of integration** of many electronic resource discovery systems, have raised the entry threshold of information technology literacy. This acts as a barrier to information search and retrieval. [...]

Users find database structures hinder. They have to learn the procedural knowledge for using a particular database as well as have some basic knowledge of how the data table is organised and what subject matter the built-in thesauri refers to; both have limited transferability. The participants did not appear to lack information technology or digital literacy, as they had demonstrated they were able to use other internet-based search and retrieval tools.“

(Wong, W. ; Stelmaszewska, H. ; Barn, B. ; Bhimani, N. ; Barn, S.: JISC User Behaviour Observational Study: User Behaviour in Resource Discovery. Final Report / JISC. Version: November 2009.
<http://www.jisc.ac.uk/media/documents/publications/programme/2010/ubirdfinalreport.pdf>)

Konsequenzen

Idee: Nur noch ein Zugangspunkt für alle Bestände (und Dienste ?!?) einer Bibliothek im Web

- große, aggregierte / integrierte Suchindexe (die alle Bestände einer Bibliothek abbilden: Monographien, Artikel, elektronische Bestände, Webseiten, ...)
- „einfach“ zu benutzende Oberflächen
- fehlt oft: gute Integration aller angebotenen Dienste (Ausleihe, Fernleihe, elektronischer Zugang,...)

Konsequenzen

Endlich Abkehr vom Paradigma exakter, boolescher Suche: „The conventional Boolean retrieval methodology is not well adapted to the information retrieval task“ (Gerard Salton, 1984)

Stattdessen einfache, „best match“ Suche („wie im Web“): „A best match search matches the set of query stems against the sets of stems corresponding to each of the documents in the database, calculates a measure of similarity between the query and each document, and then **sorts the documents into order of decreasing similarity with the query.**“ (Peter Willett, 1988)

→ „unpräzise“ Suche (egal ob [Nachname, Vorname] oder [Vorname Nachname], Stemming, nur ein Teil der Suchbegriffe muss vorkommen, automatische Erweiterung der Suchanfrage, ...)

→ Relevanzsortierung (→ was ist Relevanz?)

→ Möglichkeiten zum Browsen, Verfeinern, Entdecken (Facetten, „ähnliche Treffer“, Vorschläge/Recommendations), ...

Suchtechnik: Oft basierend auf Vektorraummodell (→ Termstastistik als Grundlage der Sortierung), z.B. Lucene/Solr/ElasticSearch/..., Xapian, ...

Salton, Gerard (1984): The Use of Extended Boolean Logic in Information Retrieval, in: Proceedings of the 1984 ACM SIGMOD International Conference on Management of Data, ACM Press, 277.

Willett, Peter (1988): Document retrieval systems, Vol. 3, Foundations of Information Science, Taylor Graham.

Aggregierte Suchindexe

Kommerzielle Angebote: Serials Solutions Summon, ExLibris Primo Central, EBSCO Discovery Service, WorldCat local metasearch, ...

Projekte/DIY: Suchkiste (Nationallizenzen), Swissbib (schweizer Bibliotheken), ETH Knowledge Portal (gesamter Bestand der Bibliothek plus Webseiten), The European Library (europäische Nationalbibliotheken), Hathi Trust (US-amerikanische Digitalisierungsprojekte), ...



Zentraler GBV Index

Inhalt:

GVK (33 Mio. Titel)

ÖVK (6 Mio. Titel)

Online Contents OLC (33 Mio. Aufsätze aus ca. 27000 Zeitschriften seit 1993)

Fachausschnitte OLC

Nationallizenzen (23 Mio. Aufsätze, 0,4 Mio. E-Books)

Bibliografien: BMS, WAO, u.w.

nationale Projekte: VD17, IKAR, usw.

Zentraler GBV Index

Suchmaschine: Solr/Lucene (Open Source):
<http://lucene.apache.org/solr/>

Indexierung von MARC21-Daten aus CBS mittels solrmarc
(Open Source): <http://code.google.com/p/solrmarc/>

Suche über „Solr-Protokoll“ (→ beliebige Oberflächen mit Schnittstellen zu Solr nutzbar, z.B. OCLC Touchpoint, VuFind, ExLibris Primo, Blacklight...)

Support der Verbundzentrale für OCLC Touchpoint, VuFind
Flexible Sichten, individuelle Ausschnitte (bibliotheks- oder nutzerspezifisch, lizenzspezifisch, thematisch, ...) auf Index mittels Filter

Derzeit Testbetrieb in zwei Installationen

TUBfind
Suchportal der Universitätsbibliothek der TU Hamburg-Harburg

Erweitert
Katalog Weblog Webseite TUBdok GBV Central TUBfind Blog
Suchhistory TagCloud

Alter Katalog

vufind
Powered by VuFind

TUBfind, TU Hamburg-Harbug (VuFind)

Bauhaus-Universität Weimar
Universitätsbibliothek

Suche Merkliste Konto Neuerwerbungswunsch Anmelden

Sprachauswahl deutsch | englisch

Willkommen

Herzlich Willkommen im Testportal Touchpoint der Universitätsbibliothek Weimar. Sie können mit einer Suchanfrage in mehreren Datenbanken recherchieren, sich ihr eigenes Rechercheprofil erstellen, Literaturlisten anlegen und bibliographische Daten in Ihr Literaturverwaltungssystem laden.

Sucheingabe

Suche eingrenzen
Sucheinstellungen
Datenbankauswahl (Bitte ANMELDEN um alle Datenbanken durchsuchen zu können)

Datenbanken
alle | keine

- Lokale Kataloge
 - Lokaler Katalog UB Weimar
- Überregionale Kataloge
 - GVK - Gemeinsamer Verbundkatalog
- Fachdatenbanken
 - Fachübergreifend
 - QLC - Online Contents
 - Elektronische Volltexte (Nationallizenzen)
 - Architektur und Kunstgeschichte
 - QLC Architektur und Kunstgeschichte
 - Kultur und Medienwissenschaften
 - QLC Kultur- und Medienwissenschaften
 - Business Source Premier
 - Bauingenieurwesen
 - QLC Bauingenieurwesen

Weitere Datenbanken:
Für die Testphase sind in Touchpoint fachbezogene Ausschnitte aus der Swets-Datenbank Online Contents sowie die Datenbank Business Source Premier enthalten. Weitere Fachdatenbanken finden Sie in unserem Datenbank-InfoSystem DRIS

Ihre Meinung ist und wichtig:
Wir haben zwei Foren (Blog und Nutzerumfrage) für Sie vorbereitet, über deren intensive Nutzung wir uns freuen würden. Lob und Tadel, aber auch Anregungen für die Weiterentwicklung des Testportals sind sehr willkommen!

Copyright © 2010. Alle Rechte vorbehalten. OCLC

TouchPoint UB Weimar

VZG

Problemchen

- Datenhandling (Normalisierung, Austausch, Anreicherung, Deduplizierung,...)
- Akzeptanz: „In both cases, the recall-oriented Solr has proven to be a very disruptive indexing technology for librarians accustomed to traditional indexes.“ (Garrison, Boston, Bair 2011)
- Integration bestehender Infrastruktur (Dienste!)
→ Weiterdenken: Discovery findet eh „überall“ statt.
Angebot von Bibliotheken ist der Zugang zu entdeckten Ressourcen. Also Zugang statt „Suche“ optimieren?

Garrison, Scott; Boston, George; Bair, Sheila (2011): Taming Lightning in More Than One Bottle: Implementing a Local Next-Generation Catalog Versus a Hosted Web-Scale Discovery Service, in: Proceedings of the ACRL 2011 Conference,
http://www.ala.org/ala/mgrps/divs/acrl/events/national/2011/papers/taming_lightning.pdf



Foto von <http://de.wikipedia.org/wiki/Datei:Fleischwolf.ganz.jpg>

Till Kinstler
kinstler@gbv.de
<http://twitter.com/tillk>