



# Relevanzranking als Erfolgsfaktor für Discoveryssysteme

Elmar Haake  
Staats- und Universitätsbibliothek Bremen

GBV-Verbundkonferenz Hannover 05.09.2012

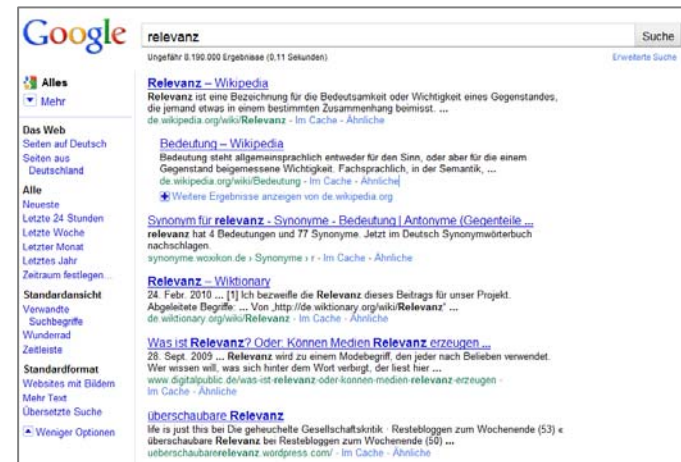
## Relevanzranking als Erfolgsfaktor für Discoveryssysteme

- Definition: Was ist Relevanzranking?
- Ausgangslage: Nutzergewohnheiten, Rankingverfahren in klassischen Bibliothekssystemen und Discoverysystemen
- Neue Lösungsansätze beim Ranking in der E-LIB Bremen



# // Definition Relevanzranking

- Das Beste für mich zuerst! Relevante Titel bevorzugen!
- Suchmaschinen: Textstatistik & linktopologisch (PageRank: Popularität der Webseiten), Boostfaktoren und nutzerspezifisches Ranking
- Nutzergerechte Sortierung der Suchergebnisse in der Reihenfolge absteigender Relevanz bezogen auf die Sucheingabe.
- Relevanz ist stets subjektiv, geringe Transparenz





## Klassische Bibliothekskataloge:

- Expertenstrategie (Exact-Match, Boolesche Logik)
- weist nur einen Teil des Bibliotheksbestandes nach
- Chronologische Sortierung als Standard
  - **Ranking nach Zeit** als relevantes Kriterium
  - bei Known-Item-Suchanfragen hilfreich
  - problematisch bei einfachen Suchanfragen bzw. hochfrequente Begriffen (große Treffermengen)
  - Zeitschriften schwer auffindbar ohne Expertenretrieval
- Verschärfung der Sortierproblematik bei Ausweitung des Suchraumes
- „Mischen impossible“?

**Better**

University of  
at typing his  
the library  
in a book about

night. At the  
ers from a

online library  
arch results by  
can befuddle  
Bauer, a

graduate student... such a hard time finding materials that she titled a bibliography "Meager Fruits of an Ongoing Fight With Virgo."



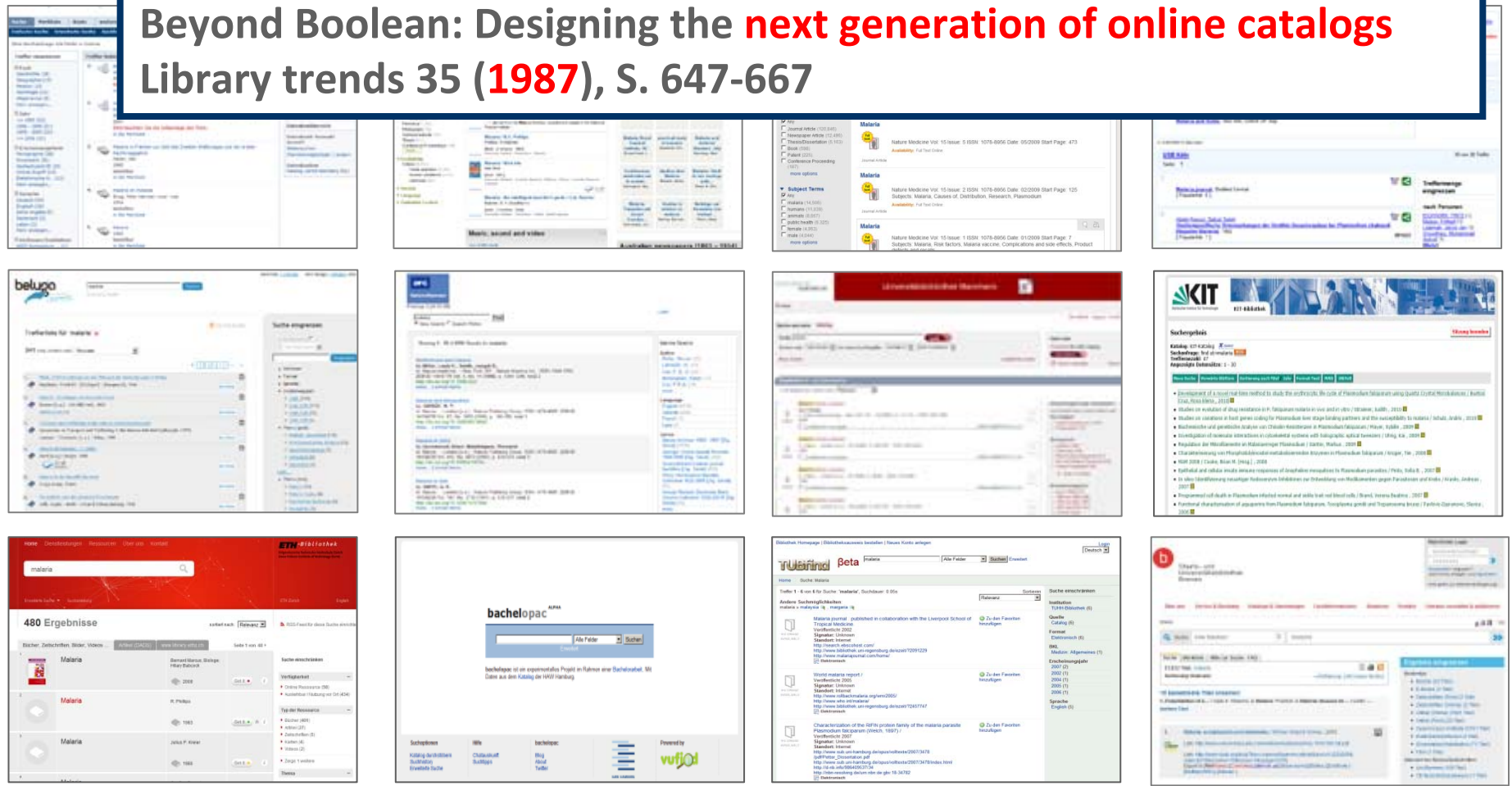
# // Neue Online-Kataloge in wiss. Bibliotheken

● Next Generation Catalogue

● Katalog 2.0

● Discovery System

Hildreth, Charles R.:  
Beyond Boolean: Designing the **next generation of online catalogs**  
Library trends 35 (1987), S. 647-667



sixties fashion (Treffer: 3)  
 med 109.3? (Treffer: 177)  
 bezugspflege (Treffer: 134)  
 mode weiblichkeit (Treffer: 31)  
 zoltai, stout: mineralogy, concepts and principles (zs) (Treffer: 0)  
 zufriedenheit schule (Treffer: 49)  
 huysen scherpe (Treffer: 4)  
 putnis: introduction to mineral science. cambridge univ. press (p) (Treffer: 0)  
 rueckstellungen bilmog (Einschränkung auf Medientyp bei der Eingabe) (Drilldown auf Zeitraum) (Treffer: 14)  
**brueckner, anthony. 1992. if i am a brain in a vat, then i am not a brain in a vat. mind 101:123-128 (Treffer: 1)**  
 hahn: international tables for crystallography. kluwer (it) (Treffer: 8)  
 pestizig (Einschränkung auf Medientyp bei der Eingabe) (Treffer: 0)  
 pestizide (Einschränkung auf Medientyp bei der Eingabe) (Treffer: 192)  
 paus-haase (Treffer: 16)

- möchten Inhalte online sofort nutzen
- finden in Bibliothekskatalogen oft nicht das passende Material

Zitierte  
 Standardformat  
 Websites mit Bildern  
 Mehr Text  
 Übersetzte Suche  
 Weniger Optionen

**Was ist Relevanz? Oder: können Medien Relevanz erzeugen...**  
 26. Sept. 2009 ... Relevanz wird zu einem Modetbegriff, den jeder nach Belieben verwendet.  
 Wer wissen will, was sich hinter dem Wort verbirgt, der liest hier ...  
[www.digitalpublic.de/was-ist-relevanz-oder-kann-medien-relevanz-erzeugen](http://www.digitalpublic.de/was-ist-relevanz-oder-kann-medien-relevanz-erzeugen)  
 Im Cache · Ähnliche

**Überschaubare Relevanz**  
 Me is just this bei Die geheuchelte Gesellschaftskritik · Restebloggen zum Wochenende (53) ·  
 überschaubare Relevanz bei Restebloggen zum Wochenende (50) ...  
[überschaubarerelevanz.wordpress.com/](http://überschaubarerelevanz.wordpress.com/) · Im Cache · Ähnliche

## // Eigenschaften von Discoverysystemen:

- Großer Suchraum - Große Treffermengen
- Best-Match-Verfahren
- Relevanzranking über textstatistische Verfahren (u.a. Gewichtung von Termen im Dokument)
- Relevanzdifferenzierung schwierig bei Metadaten , großen Treffermengen und hochfrequenten Begriffen
- Facettierung als ‚Notlösung‘, aufwändige Implementierung, Usability problematisch
- Differenzierung bei weiterhin identischen bzw. eng liegenden Relevanzwerten?
  - Chronologische Sortierung?
  - Zufällige Sortierung?
  - Boosting durch Bevorzugung bestimmter Titel (Materialtyp : Online-Ressourcen, Zeitschriften, Lehrbücher oder lokaler Bestand)?





Wie kann das Ranking nutzergerecht weiter verbessert werden?





## // Relevanzranking: Modifikation über Titelpopularität in der E-LIB Bremen

Hilfsmittel: Suchhistorie der Nutzer in anonymisierter Form, Erfüllung der Nutzererwartungen analysieren

Zieldefinition:

Bevorzugung stark nachgefragter Titel bei gleicher Relevanz und ‚dezentere‘ Platzierung der Artikel mit vertretbarem Aufwand und möglichst geringer Komplexizität

Lösungsansatz:

Zusätzliche Beeinflussung der Relevanzparameter durch Titelpopularität\* als Ergänzung der textstatistischen Verfahren

- Beschaffungsverhalten
- Medieneigenschaften
- Nutzerverhalten

\* D. Lewandowski, Ranking library materials. Library Hi Tech 27 (2009) 4, 584 - 593 - <http://kurl.de/elib-lw>





## // Relevanzranking: Modifikation über Titelpopularität in der E-LIB Bremen

### Beschaffungsverhalten

- Stark gefragte Titel werden häufig als Staffelexemplare von der Bibliothek beschafft:  
Exemplarzahl > 3 Exemplare als Indiz für Popularität,  
Relevanzaufschlag





## // Relevanzranking: Modifikation über Titelpopularität in der E-LIB Bremen

### Medieneigenschaften

- Stark gefragte Titel sind häufig im Buchhandel vergriffen und werden als Folgeauflagen neu aufgelegt:  
Auflagenzählung als weiteres Indiz für Popularität, Relevanzaufschlag





## // Relevanzranking: Modifikation über Titelpopularität in der E-LIB Bremen

### Nutzerverhalten

- Stark gefragte Titel werden in der Trefferliste häufig angeklickt
- Aktualität der Titel relevant, daher Auswertung der Klicks auf die letzten 2 Jahre beschränkt
- Klicks > 3 als weiteres Indiz für Popularität, Relevanzaufschlag





## // Relevanzranking: Modifikation über Titelpopularität in der E-LIB Bremen

Keine Überkompensation der Textstatistik durch Titelpopularität

Grundlage: Textstatistische Berechnung der Relevanzwerte und Feinjustierung durch Nutzung von Popularitätsfaktoren

Überblick:

<b>Popularitätstyp</b>	<b>Relevanzaufschlag in %</b>
Beschaffungsverhalten	3
Medieneigenschaften	3
Nutzerverhalten	5





## // Relevanzranking: Modifikation über Titelpopularität in der E-LIB Bremen

### Weitere mögliche Beeinflussungsparameter

Popularität :

- Ausleihstatistik (beeinflusst nur ausleihbaren Bestand, keine Online-Ressourcen)
- Zitationsfrequenz (problematische Beschaffung des Datenmaterials)
- Nutzerbewertung (ausgeprägte Subjektivität?)
- Renommee der Autoren (Objektivierbarkeit?)
- Impact Factor der Zeitschrift bei Zeitschriftenartikel (Datenmaterialbeschaffung?)

Lokalität: eigenen Bestand bevorzugen, standortspezifische Sortierung

Medientyp : Lehrbücher für UB, Aufsätze für Virtuelle Fachportale, bei homogenem Nutzertyp möglich

Relevanzranking als Erfolgsfaktor:

- Discoverysysteme können für die Mehrzahl der Nutzer mit Textstatistik besser als klassische Kataloge ranken.
- Bei großen Treffermengen besteht jedoch ein weiterer Justierungsbedarf, der möglichst im Hinblick auf die Erfüllung der Nutzererwartungen mit Hilfe von Popularitätsfaktoren eingestellt werden kann.
- Kontinuierlicher Überprüfungsbedarf, unterschiedliche Nutzererwartungen, keine vollständige Objektivierbarkeit, schwierige Evaluierbarkeit

**Mutig sein  
und probieren**





# Relevanzranking als Erfolgsfaktor für Discoverysysteme

Elmar Haake  
Staats- und Universitätsbibliothek Bremen

GBV-Verbundkonferenz Hannover 05.09.2012