

Wie können sich Bibliotheken gegenüber Wissenschaftssuchmaschinen positionieren?

Dr. Dirk Lewandowski
dirk.lewandowski@uni-duesseldorf.de

www.durchdenken.de/lewandowski

Gliederung

Web-Trends

Das Academic Web

Überblick Wissenschaftssuchmaschinen

Chancen und Herausforderungen für Bibliotheken

Gliederung

Web-Trends

Das Academic Web

Überblick Wissenschaftssuchmaschinen

Chancen und Herausforderungen für Bibliotheken

Vom Anbieter zum Nutzer: Das „Mitmach-Web“ (Web 2.0)

Web-Trends

- **Kollaboratives Erarbeiten von Inhalten, Netzwerkbildung, „von Nutzer zu Nutzer“.**
 - Wikis (Bsp. Wikipedia)
 - Nutzerbewertungen: (Bsp. Amazon, Ebay, Kelkoo)
 - Frage-und-Antwort-Dienste (Bsp. Lycos IQ)
 - Social Networking (Bsp. MySpace, OpenBC),
- **Jeder ist selber Journalist.**
 - Blogs
- **„Tagging“ ersetzt Verschlagwortung.**
 - Inhalte werden mittels freier Wörter von den Nutzern erschlossen
 - Beispiele: Flickr.com, del.icio.us
 - Stichwort „Folksonomies“

» del.icio.us history for

Inhaltsverzeichnis aus: Dr. Dirk Lewandowski: Web Information Retrieval

<http://www.durchdenken.de/lewandowski/web-ir/>

this url has been saved by 33 people.
[save this item to your bookmarks »](#)

user notes

Mar '06

Dirk Lewandowskis „Web Information Retrieval“ ist eine methodisch fundierte informationswissenschaftliche Bestandsaufnahme der Web-Suchmaschinen. Schon dies würde die Forschung weiterbringen, da eine solche deskriptiv-analytische Zusammenschau bislang

- [gordimer](#)

buch online

- [krueggen](#)

Feb '06

Inhaltsverzeichnis des eBooks: Web Information Retrieval. Technologien zur Informationssuche im Internet.

- [wibiboka](#)

Jan '06

Dirk Lewandowski stellt Technologien zur Informationssuche im Internet vor

- [any.user](#)

Dec '05

Web Information Retrieval Technologien zur Informationssuche im Internet Dirk Lewandowski Taschenbuch - 248 Seiten - Deutsche Gesellschaft f. Informationswissenschaft u. Informationspraxis - 2005 ISBN: 3925474552 - 25,00 Euro (Mitglieder der DGI erhal

common tags [cloud](#) | [list](#)

- 8 [web](#)
- 7 [informationswissenschaft](#)
- 6 [searchengine](#)
- 5 [information_retrieval](#)
- 4 [recherche](#)
- 4 [retrieval](#)
- 4 [search](#)
- 3 [book](#)
- 3 [information](#)
- 3 [informationsmanagement](#)
- 3 [suchmaschine](#)
- 3 [suchmaschinen](#)
- 2 [ebook](#)
- 2 [german](#)
- 2 [informationretrieval](#)
- 2 [science](#)
- 2 [seo](#)
- 2 [studium](#)
- 2 [tutorial](#)

related items - [show ↓](#)

posting history

» first posted by [slingshot](#) to [bibliothekswissenschaft library](#)
[search science ebook informationsbroker studium web](#)

Mar '06



- E-Mail & Chat
- Suche
- Life@Lycos
- Webhosting
- Entertainment
- Shopping

Webkatalog Starsuche Profisuche Livesuche Popularimeter Lycos Slim Search Seite anmelden

Web | [News](#) | [Bilder](#) | [Produkte](#)

Finde: [Erweiterte Suche](#) [Einstellungen](#) [Frage jetzt die Experten von LYCOS iQ](#)

Suche: Das Web Seiten aus Deutschland

Web: Suchergebnisse für "vogelgrippe"

Sponsored Links:

- [Schutz bei Vogelgrippe für Kinder und Erwachsene](#) - Zertifizierte Schutzmaskenprodukte - informieren und online kaufen. www.pandemie.com
- [Vogelgrippe, H5N1 - medicalpicture zeigt Hintergründe](#) - 3D - Grafiken, Bilder und Fotos des H5N1 Virus, Menschen mit Mundschutz, Influenza A Virus, Haushühner, Infografiken und Illustrationen zur Hühnerpest. Neue Bilder zum Aufbau des Virus. Wie ist das Virus ... vogelgrippe.medicalpicture.de

Lycos iQ Fragen & Antworten - [Zeige alle Fragen & Antworten](#)

Frage: Wo finde ich Informationen über den "Fortschritt" der Vogelgrippe in anderen Ländern?
Antwort:H5N1-Virus bislang in 25 Ländern nachgewiesen Der Virus-Tod von Schwänen und Wildvögeln in Deutschland gibt erste Hinweise darauf, dass die hier angekommene ...
 Gespeichert in: [deutschland Vogelgrippe EU](#) von [Momo23](#) am 24.Feb.2006

Frage: Was wird mich die Vogelgrippe kosten?

Antwort:du kannst daran sterben ...
 Gespeichert in: [Vogelgrippe](#) von [hessej](#) am 15.Feb.2006

Lycos iQ Experten Links - [Zeige alle Experten Links](#)

Influenza, Vogelgrippe und Pandemie: Drohende Gefahr? - Ratgeber | SWR.de
 Infos zur Vogelgrippe
 Gespeichert in: [deutschland Gefahr Vogelgrippe Geflügel Pandemie Influenza](#) von [yoyo4](#)

Die TOP 10 Deutsche Seiten von 808.559

[Vogelgrippe - Avian-Flu](#)

Vogelgrippe. (Geflügelpest, Aviäre Influenza A/H5(N1), Avian Flu, Highly Pathogenic Avian Influenza HPAI) Stand: 16. Februar 2006. Zusammenfassung ...

Lycos Topsuchen:

- Privat Krankenversicherung
- Geldanlage
- Immobilie
- Autoteil
- Buch

Lycos iQ Search

Verwandte Themen:

- [deutschland](#)
- [tiere](#)
- [virus](#)
- [influenza](#)

Experten zum Thema:

- [daw2craw](#)
Rank: Student
- [Weisser.Scorpion](#)
Rank: Igor Sikorsky

[vogelgrippe](#)
 eBay Angebote zum Thema vogelgrippe
HIER KLICKEN!
www.ebay.de

[Der Vogelgrippe vorbeugen mit](#)

Die Suche wird zum zentralen Konzept.

Allgemeine Trends in der Suche

- **Benutzerführung *nach* der Präsentation der ersten Ergebnisse.**
 - Vorschläge zur Erweiterung/Eingrenzung der Suche
 - Verweise auf Spezialsuchen (News, Produkte, Bilder, usw.)
- **Personalisierung**
 - Ergebnisse (hier auch: Pushdienste, Alerts)
 - Ergebnispräsentation
- **Suche findet nicht mehr nur in Suchmaschinen statt.**
 - „Ambient findability“: Desktop, Mobiltelefon, andere mobile Endgeräte, ...
- **Suchmaschinen werden verstärkt zu „Trefferlieferanten“.**
 - „White Label“ Suchlösungen (auf Portalen: T-Online, Web.de, usw.)

Bereiche des (Academic) Web

- **Surface Web**

- Alle Inhalte, die von den allgemeinen Suchmaschinen erschlossen werden (können).

- **Invisible Web**

- Alle Inhalte, die von den allgemeinen Suchmaschinen nicht erschlossen werden (können), vor allem die Inhalte von Datenbanken, die über das Web erreichbar sind.

- **Academic Surface Web**

- Wissenschaftliche Inhalte im Oberflächenweb.
- Alle Seiten von Unis, Forschungseinrichtungen, usw.
- Wissenschaftliche Texte.

- **Academic Invisible Web**

- Vor allem Inhalte aus Datenbanken.
- Umfang: Ähnliche Größe wie Datenbestände der allgemeinen Suchmaschinen.

Wissenschaftssuchmaschinen werden das Academic Web erschließen.

Die Inhalte der Wissenschaftssuchmaschinen

- **Verlagsinhalte**
 - Bücher
 - Aufsätze
- **Graue Literatur aus dem Web**
 - Reports
 - Manuskripte
- **Open Access**
 - Zeitschriften
 - Repositorien
- **Inhalte aus Datenbanken**
- **(Forschungsdaten)**

Gliederung

Web-Trends

Das Academic Web

Überblick Wissenschaftssuchmaschinen

Chancen und Herausforderungen für Bibliotheken

Google Scholar

- **Suchmaschine für wissenschaftliche *Literatur***
 - Aufsätze
 - Bücher
- **Gewisser Anspruch auf Vollständigkeit**
 - Volltexte + Zitationen
 - „Auf die Texte, die wir nicht selbst im Bestand haben, verweisen wir...“
- **Zitationsanalyse**

Motto: Alle Fächer, alle Quellen, alle Texte

Inhalte von Google Scholar

- **Fächer**

- prinzipiell alle Fächer, bisheriger (beobachteter) Schwerpunkt bei Naturwissenschaft und Technik (v.a. aufgrund der Struktur des Web und der Verlagslandschaft).

- **Quellen**

- freies Web
- Verlage und Fachgesellschaften
- Open-Access-Archive
- **kein** Quellenverzeichnis; Umfang der Quellen unklar

- **Sprachen**

- prinzipiell alle

Inhalte werden gecrawlt und im Volltext erschlossen.

Einbindung und Erschließung der Inhalte

- **Erfassung der Inhalte**

- crawlerbasiert; keine Feeds von Verlagen
- Crawling von geschützten Inhalten bei den Verlagen

- **Kooperationen**

- keine veröffentlichte Liste der Kooperationspartner
- u.a. ACM, Blackwell, Institute of Physics, Nature Publishing Group, Wiley, Springer

- **Erschließung**

- Volltexterschließung (+Relevance Ranking)
- keine Übernahme der vorgegebenen Schlagwörter, Systemstellen, usw.
- keine Anwendung linguistischer Verfahren
- **Die Erschließung der Dokumente in Feldern ist absolut unzuverlässig.**

Windows Live Academic

- **Suche in Verlagsarchiven (Aufsätze).**
- **Bereits in die MSN-Suche eingebunden.**
- **Frühes Entwicklungsstadium.**

Andere

Andere

- **Scirus**

- Academic Surface Web (keine Beschränkung auf Literatur)
- Teile des Academic Invisible Web
- Elsevier-Content

- **Forschungsportal.net**

- Websites der in Deutschland öffentlich geförderten Forschungseinrichtungen
- Online-Dissertationen DDB

- **Nicht zu vergessen:**

- Interdisziplinäre Literaturdatenbanken (Web of Science, Scopus)
- Verlagsangebote (Springerlink, Science direct).

Google Buchsuche

- **Suche in digitalisierten Büchern, echte Massendigitalisierung.**
- **Volltexte durchsuchbar, bei gemeinfreien Werken Download als PDF möglich.**
- **Alle Werke sind durchsuchbar und zumindest in Ausschnitten ansehbar.**
- **Keine weitere Erschließung.**
- **Qualität der Scans teils schlecht.**

Open Content Alliance: Das „Gegenmodell“ zur Google Buchsuche

- **Enthält nur freie Werke und von den Verlagen explizit genehmigte.**
- **Kooperationspartner u.a.: Web Archive (archive.org), Yahoo, MSN.**
- **Inhalte werden vollständig frei verfügbar gemacht.**

Amazon: Im Idealfall umfassende Informationen zum Titel.

- **Informationen u.a.:**

- Bibliographische Angaben
- klassifikatorische Angaben
- Schlagwörter
- Klappentext
- Besprechungen („Redaktion“ + Kunden)
- Hinweise auf ähnlich Bücher
 - aufgrund des Kaufverhaltens
 - aufgrund des Browsingverhaltens
- wichtige Mehrwortausdrücke aus dem Text
- Zitationen
- Von Kunden vergebene *Tags*
- von Kunden erstellte Themenlisten
- beschränkt zugänglicher Volltext („Search Inside“)
- „Upgrade“: Zusätzlich zum gedruckten Buch die elektronische Version mit der Möglichkeit der Bearbeitung.

- **Beispiel**

http://www.amazon.com/gp/product/1558607544/sr=8-1/qid=1144187327/ref=sr_1_1/002-9049787-9343215?%5Fencoding=UTF8

Suchmaschinen dringen in weitere Suchbereiche der Bibliotheken vor.

- **Google News Archive**

- Ähnliches Konzept wie Google Scholar.
- Kostenlose und kostenpflichtige Inhalte.

- **Yahoo Subscriptions**

- Durchsucht optional Factiva, LexisNexis, IEEE, Forrester Research, u.a. zusammen mit der Websuche.
- Lizenz für die Quellen notwendig.

→ Diese Dienste sind problemlos auf andere Bereiche übertragbar.

Der Anspruch der Wissenschaftssuchmaschinen geht weit über den der Bibliothekskataloge hinaus.

Vorteile der kommerziellen Wissenschaftssuchmaschinen

- **Inhalte**

- Alle Aufsätze, alle Bücher.
- (teils) andere Inhalte des Academic Web.

- **Erschließung**

- Volltexterschließung.
- Anreicherung durch Volltexte (bzw. Ausschnitte), Rezensionen, „tags“
- Empfehlungssysteme.

- **Suche**

- Schnelle und einfache Suche.
- Suchinterfaces wie bei allgemeinen Suchmaschinen.

Wissenschaftssuchmaschinen zeigen wenig Transparenz, die Inhalte sind meist schlecht erschlossen.

Nachteile der kommerziellen Wissenschaftssuchmaschinen

- **Unklare Quellenlage**
 - Ausgewertete Quellen werden nicht angegeben.
 - Quellen werden oft nicht vollständig erschlossen.
- **Mangelhafte Erschließung**
 - Hohe Fehlerrate (Autorennamen, Zeitschriftentitel)
 - Keine Erschließung mit Klassifikation, Schlagwörtern, usw.
- **Zu allgemeine Community**
 - Rezensionen bei Amazon werden von *irgendwelchen* Nutzern geschrieben.

Bibliotheksangebote sind bisher nur Nachahmer, nicht Vorreiter.

Wissenschaftssuchmaschinen aus der Bibliothekswelt

- **BASE**
 - Teilbestand des OPAC UB Bielefeld.
 - ca. 160 Open-Access-Quellen.
- **HBZ-Suchmaschine**
 - OPAC auf Basis von Suchmaschinen-Technologie.
 - Technologie kann aufgrund des geringen Textumfangs der Datensätze nicht vollständig ausgenutzt werden.
- **Vascoda**
 - Meta-Ansatz; schrittweise Umstellung auf Suchmaschinentechnologie.
- **Dandelon.com**
 - Anreicherung von Katalogdaten durch Inhaltsverzeichnisse.
 - Automatische Indexierung.

Bibliotheken als Innovatoren

- **Technische Ebene**

- Zur Zeit Ablösung der alten DB-Technologie durch Suchmaschinen-Technologie.
- Technologieentwicklung durch Bibliotheken aussichtslos; nur in Zusammenarbeit mit starken Partnern.
- Hersteller von Bibliothekssystemen müssen diesen Wandel schnellstmöglich vollziehen.

- **Benutzerebene**

- Wenn sich Bibliotheken konsequent an den Nutzerbedürfnissen orientieren und entsprechende Informationssysteme entwickeln, können sie hier Vorreiter (auch für andere Bereiche) sein.
- Stichwort „Benutzerführende Systeme“.

- **Erschließungsebene**

- Stärken der Bibliotheken in die zukunftsorientierten Anwendungen „hinüberretten“.

Stärken der kommerziellen Angebote adaptieren und durch bibliothekarische Stärken erweitern.

- **Inhalte**

- Erweiterung der Kataloge um Aufsätze, Web-Inhalte.
- Dem Nutzer jederzeit klarmachen, wer die Inhalte bereitstellt.

- **Anreicherung der Titelaufnahmen**

- Durchsuchbare Inhaltsverzeichnisse (und Volltexte?).
- Klappentexte.
- Zugekaufte Rezensionen u.ä.

- **Nutzergenerierte Inhalte**

- Rezensionen von vertrauenswürdigen Bibliotheksnutzern.
- Rezensionen durch unterschiedliche Nutzergruppen (Studenten, wissenschaftliches Personal).

- **Frage- und Antwort- Dienste**

- Vertrauenswürdige Antworten aus dem (elektronischen) Auskunftsdienst.
- Durchsuchbares Archiv der Antworten.

- **Empfehlungssysteme**

- Automatisch generierte Empfehlungen aufgrund des Ausleihverhaltens.

- **Erschließung**

- Alte Stärken nutzen und besser herausstellen.

Vielen Dank für Ihre
Aufmerksamkeit.

www.durchdenken.de/lewandowski

E-Mail: dirk.lewandowski@uni-duesseldorf.de