

**Zeitschrift:** Arbido  
**Herausgeber:** Verein Schweizerischer Archivarinnen und Archivare; Bibliothek Information Schweiz  
**Band:** - (2008)  
**Heft:** 4: Informationswissenschaft: die Instrumente der Zukunft = Information documentaire: les outils du futur = Scienze della informazione: gli strumenti di domani  
**Rubrik:** Recensions = Buchbesprechungen

### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

### **Conditions d'utilisation**

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

### **Terms of use**

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

**Download PDF:** 09.01.2026

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

### III. Recensions

## Buchbesprechungen

# Schlag nach bei Lewandowski

Stephan Holländer

**Mit einem Handbuch zum Thema Suchmaschinen legt Dirk Lewandowski ein konzeptuell gelungenes Werk vor<sup>1</sup>. Die folgende Detailkritik soll als Plädoyer für ein eingehendes Fachlektorat in einer künftigen Auflage verstanden werden. Trotzdem lohnt sich die Anschaffung: Eine anregende Auseinandersetzung mit dem Thema ist garantiert.**

Der Herausgeber und Mitautor Dirk Lewandowski legte zum Bücherherbst an der Frankfurter Buchmesse ein Buch vor, das einen Überblick über die aktuellen Tendenzen bei Suchmaschinen für das World Wide Web liefert. Wie der Herausgeber in seinem Vorwort schreibt, sollen Beiträge von Wissenschaftlern und Praktikern den Entwicklungsstand und die Tendenzen in dem jeweiligen Gebiet aufzeigen. Die *Nutzerorientierung in Wissenschaft und Praxis*, so der Untertitel, ist das Leitmotiv der publizierten Beiträge im Buch.

Der Bereich Web Information Retrieval ist ein dynamisches Forschungsfeld, das eine veritable Publikationsflut von Forschungsergebnissen hervorgebracht hat. Man spürt die Intention des Herausgebers, mit der Publikation zu zeigen, dass sich die klassische Informationswissenschaft neuen Gebieten öffnen muss, um die Tendenzen des Web Information Retrieval weiter erforschen zu können. Eine grosse Herausforderung ergibt sich angesichts der Tatsache, dass ein grosser Teil der intern in Unternehmen erzielten For-

schungsergebnisse nicht veröffentlicht wird, um dem eigenen Unternehmen Konkurrenzvorteile gegenüber anderen Unternehmen der Branche zu verschaffen. Ein Ziel des Herausgebers ist es sicher, einer verstärkten Zusammenarbeit zwischen akademischer Forschung und Suchmaschinen-Unternehmen den Weg zu bahnen.

Der Herausgeber verspricht uns ein Handbuch. An ein Handbuch dürfen ganz spezifische Erwartungen gestellt werden. Der Brockhaus<sup>2</sup> meint dazu trocken: «Werk, das ein komplexes Fach- oder Wissensgebiet übersichtlich darstellt.» Das Wörterbuch des Buches<sup>3</sup> führt zum Stichwort etwas eloquenter aus: «Bezeichnung für ein Buch von <handlichem> Umfang und Format, das in der Regel den Stoff eines Wissensgebietes in kurz gefasster, aber doch eingehender Form bietet.» In anderen Werken wird auch ein Glossar als formbestimmendes Element genannt. Im Folgenden sollen einige ausgewählte Kapitel aus Praktikersicht näher betrachtet werden:

**Christian Maass, Andrea Skusa, Andreas Hess, «Der Markt für Internetsuchmaschinen»**

Die Autoren legen mit ihrem Beitrag eine gute Auslegeordnung des Suchmaschinen-Marktes vor. Der Beitrag bietet eine kompakte Übersicht über die Rahmenbedingungen und das Marktvolumen des Suchmaschinen-Marktes. Leider wird das positive Bild, das man bei der Lektüre gewinnt, durch kleine Fehler bei den Fakten getrübt.

Hier – wie auch an einigen anderen Stellen im Buch – wäre ein konsequentes Fachlektorat angebracht gewesen. So wird beispielsweise das CERN als schweizerisches Kernforschungsinstitut bezeichnet<sup>4</sup> (S.4). Da freut sich zwar der Schweizer Leser, aber ein Blick auf die Homepage des CERN genügt, um festzustellen, dass 20 Länder zu den finanziellen Trägern des CERN gehören.<sup>5</sup>

**Joachim Griesbaum, Bernhard Bekavac, Marc Rittberger, «Typologie der Suchdienste im Internet»**

Mit diesem Kapitel wird ein sehr gut strukturierter Beitrag zu den verschiedenen Typen von Suchmaschinen publiziert. Die Autoren unterscheiden in ihrer Darstellung zwischen manuell und automatisch erstellten Dokumentensammlungen. In einem weiteren Schritt werden die Funktionsweisen der beiden Ansätze erklärt. Die Verfasser gehen dabei auch auf Spezialsuchdienste und Metasuchmaschinen ein. Interessant für den Leser an diesem Beitrag ist, dass die vier wesentlichen Faktoren, die das Ranking der Resulta-

- 1 Dirk, Lewandowski (Hrsg.), *Handbuch Internet-Suchmaschinen. Nutzerorientierung in Wissenschaft und Praxis*, Akademische Verlagsgesellschaft AKA, Heidelberg, 2009.
- 2 *Ausgabe 2006 in 30 Bänden*, Mannheim 2006.
- 3 Helmut Hiller/Stephan Füssel, *Das Wörterbuch des Buches*, Frankfurt 2006.
- 4 Lewandowski, S. 4.
- 5 <http://dg-rpc.web.cern.ch/dg-rpc/Scale.html>

te bei Google beeinflussen (On Page, On Site, Linkfaktoren sowie das Verhalten der Benutzer), erklärt werden. Der Beitrag zeigt, dass es einen Gewinn für ein Handbuch sein kann, wenn didak-

tische Erfahrungen in die Abfassung eines Kapitels einfließen.

#### **Dirk Lewandowski, «Spezielsuchmaschinen»**

Diese speziellen Suchangebote umfassen die Suche in breiten Themenbereichen von Nachrichten, in Blogs über spezielle Wissenschaftsbereiche bis zur lokalen Suche und der Suche nach Produkten. Die über Webcrawler gefundenen Dokumente und das Ranking der Suchergebnisse unterscheiden sich von populären Suchmaschinen im Web. Die wirtschaftlichen Bedingungen für den Betrieb von Suchmaschinen werden durch hohe Kosten für die Erstellung und Aktualisierung des Suchindex und die notwendige hohe Bandbreite bestimmt. Wie der Verfasser richtig bemerkt, sind die Rankingtechnologien noch zu sehr adressbasiert<sup>6</sup>. Der Beitrag gibt einen ausgezeichneten Überblick über den aktuellen Stand und die Tendenzen bei den Spezielsuchmaschinen. Der Beitrag weckt das Interesse an weiterführenden Informationen. Die Hinzunahme themenspezifischer Suchmaschinen in einem weiteren Kapitel wäre ein möglicher thematischer Ausbau für eine künftige Auflage des Buches.

#### **Carsten D. Schultz, «Suchmaschinenmarketing»**

Der Autor geht in einem ersten Teil auf die Suchmaschinenoptimierung (Search Engine Optimization =SEO) ein. Dies ist der Fachbegriff für Massnahmen, die dazu dienen, dass Webseiten in den Ergebnislisten von Suchmaschinen auf den vorderen Plätzen erscheinen. In einem zweiten Teil setzt sich der Verfasser mit den Aspekten der Suchmaschinenwerbung auseinander. Der Autor versucht das schier Unmögliche und stellt die beiden vorgenannten Aspekte in einem einzigen, äusserst kompakt geschriebenen Kapitel dar. Sehr positiv an dem Beitrag ist, dass es sich nicht einfach um einen Beitrag zur Suchmaschinenoptimierung bei Google handelt (zu diesem Thema gibt es ja bereits eine grosse Anzahl von Publikationen auf dem Markt). Die Quellenangabe bei der Abbildung 1 zur Aufteilung des Suchmaschinenmarkts sollte nachgeliefert werden<sup>7</sup>. Diese sehr zukunftssträchtigen Themen sollten künf-

tig in zwei getrennten Kapiteln dargestellt werden.

#### **Julian Bahrs, «Enterprise Search – Suchmaschinen für Inhalte im Unternehmen»**

Zu diesem Thema gibt es erst wenige Buchveröffentlichungen in deutscher Sprache. Gut schildert der Autor in einem ersten Abschnitt die Vision eines integrierten Informationszugangs. In einem zweiten Abschnitt wartet der Autor mit den Resultaten einer eigenen Umfrage über den Einsatz dieser Technologie bei 140 Unternehmen in der Bundesrepublik Deutschland auf. Der Verfasser lässt uns mit den Resultaten an der Erwartungshaltung der Mitarbeiter dieser Unternehmen teilhaben. Im Beitrag wird auf die niedrige Nutzerfrequenz verwiesen. Wieso die Nutzerfrequenz so niedrig bleibt, wird nicht schlüssig beantwortet. Ein Blick in andere Branchenstudien gibt jedoch Aufschluss darüber, wie es zu diesem Nutzerverhalten kommt<sup>8</sup>.

Im Abschnitt zum Funktionsmodell von Suchmaschinen wird die Behauptung aufgestellt, dass zur Informationsakquise auch die Informationsaufbereitung gehöre. Leider bleibt diese ungewöhnliche Behauptung unbelegt und wird auch nicht mit einer Literaturangabe untermauert. Die Wortwahl *Informationsaufbereitung* ist in diesem Zusammenhang unglücklich gewählt. Besser wäre der Ausdruck *Indexierung*.

Die im Beitrag angeführte Modellierungssprache KDML wurde an der Uni Potsdam entwickelt, wo der Autor tätig ist. Es steht einem Beitrag in einem Handbuch wohl an, wenn auch auf andere Methoden, die sich am Markt erfolgreich halten, hingewiesen wird<sup>9</sup>. Bei den weiterführenden Literaturhinweisen fehlen einige neuste einschlägige englischsprachige Literaturangaben<sup>10</sup>.

#### **Sonja Quirnbach, «Universal Search»**

Ein richtiger konzeptioneller Entscheid des Herausgebers ist es, ganz neue Trends in das Handbuch aufzunehmen. Das Risiko, bei solchen Beiträgen zu jungen Wein angeboten zu bekommen, trägt der Herausgeber. Der vorliegende Beitrag versucht, das Konzept der Anreicherung von Websuchresultaten mit

6 Siehe etwa <http://www.local.ch/>

7 Lewandowski, S. 70.

8 Laut einer Studie von IDC Deutschland (zitiert nach [http://www.pcwelt.de/start/computer/netzwerk\\_server/praxis/176054/volltext-suche\\_im\\_eigenen\\_netzwerk/index.html](http://www.pcwelt.de/start/computer/netzwerk_server/praxis/176054/volltext-suche_im_eigenen_netzwerk/index.html)) suchen Mitarbeiter in den von IDC befragten Unternehmen im Schnitt 15 Mal am Tag nach Informationen. Die in der Studie genannten Wissensarbeiter begeben sich sogar 30 Mal täglich auf die Suche. Das ernüchternde Fazit der genannten Untersuchung: 40 bis 50 Prozent der Suchen bleiben erfolglos. Die genannten Angestellten vergeuden also wertvolle Zeit mit der Suche nach Informationen, die sie letztlich nicht finden können.

9 So sollten auch Modellierungen wie PROMOTE Erwähnung finden, die in einem EU-Projekt entstanden ist und von der Firma BOC zu einem Produkt weiterentwickelt wurde (siehe <http://www.boc-eu.com/jumpto.jsp?goto=KNOWLEDGE%20M&lg=en>). Weiterführende Literaturhinweise: D. Karagiannis, R. Telesko: *The EU-Project PROMOTE: A Process-Oriented Approach for Knowledge Management* (2000). *Proc. of the Third Int. Conf. of Practical Aspects of Knowledge Management PAKM 2000*. Hinkelmann, K.; Karagiannis, D.; Telesko, R., PROMOTE – Methodologie und Werkzeug zum geschäftsprozessorientierten Wissensmanagement. In: Abecker, A.; Hinkelmann, K.; Maus, H.; Müller H.J. (Hrsg.): *Geschäftsprozessorientiertes Wissensmanagement*, Heidelberg, 2002. R. Woitsch, D. Karagiannis: *Process-oriented knowledge management systems based on KM-services: The promote approach* (2002). *Proceeding of the fourth International Conference on Practical Aspects of Knowledge Management (PAKM2002)*. Eine andere erwähnenswerte Methode zum Geschäftsprozessorientierten Wissensmanagement ist zum Beispiel GPO-WM von Peter Heisig (Peter Heisig: *GPO-WM: Methode und Werkzeuge zum geschäftsprozessorientierten Wissensmanagement*. In: Abecker, A.; Hinkelmann, K.; Maus, H.; Müller H.J. [Hrsg.]: *Geschäftsprozessorientiertes Wissensmanagement*, Springer-Verlag).

10 White, Martin Titel: *Making search work: implementing web, intranet and enterprise*

den Ergebnissen anderer vom Suchmaschinenanbieter angebotener Quellen (Bilder, Videos, Nachrichten) zu erklären. Wie die Autorin selber feststellt, ist das User Centered Design (UCD) eine «recht junge Disziplin» im deutschsprachigem Raum, zu der es noch «keine allgemein anerkannte Definition des Begriffs» gibt<sup>11</sup>. Dieser Befund muss nicht auf das UCD beschränkt bleiben. Bei der Begriffswahl gibt uns die Verfasserin selbst ein interessantes Beispiel. Bei der Beschreibung, wie ein Nutzer die Webseite ansieht und auf Treffer hin analysiert, gebraucht die Autorin den Begriff «Scannen»<sup>12</sup>. Der verwendete Begriff wird aber in der Regel anders definiert. Gemäss Duden ist damit der photomechanische Abtastvorgang gemeint<sup>13</sup>. Hier zeigt sich einerseits erneut die Wichtigkeit eines gründlichen Fachlektorats, damit Begriffe in einem Handbuch kohärent verwendet werden, und andererseits die Notwendigkeit zur Schaffung eines Glossars. So erhält der zu vermittelnde Inhalt Handbuchcharakter.

#### Karsten Weber, «Moral und Suchmaschinen»

Der Verfasser ist bemüht, dem allumfassenden Charakter eines Beitrages in einem Handbuch gerecht zu werden. Wie jedoch der Herausgeber in seinem Vorwort zum Handbuch schreibt, wendet sich dieses auch an Praktiker und Information Professionals, die nur zu einem kleinen Teil einen philosophischen Hintergrund haben. Diesen wird der Einstieg in die Thematik einer «libertären Informationsethik» aber alles andere als leicht gemacht, weil zu viel als bekannt vorausgesetzt wird. Zu schreiben<sup>14</sup>, dass es das Web 2.0 mit seinen Anwendungen Studi VZ oder auch YouTube den Betreibern der Angebote leicht macht, personenbezogene Daten zu speichern, mag ja gerade noch angehen. Wird dann aber im Folgenden behauptet, Google könne mit seiner Privacy Policy von einem «informed consent» der Nutzer ausgehen<sup>15</sup>, lässt der geneigte Leser das Buch erstaunt sinken. Ein Blick in die deutsche Presse belehrt uns des genauen Gegenteils<sup>16</sup>.

Für den weder ökonomie- noch philosophiebewanderten Leser ist es ein Ärgernis, wenn der Autor die Namen

von Amartya Sen und Martha Nussbaum ohne weiterführende Literaturhinweise und Erklärungen einführt<sup>17</sup>.

Hier müsste ein Fachlektorat nachhaltig eingreifen, denn es ist dem fachfremden Leser in einem Handbuch nicht zuzumuten, dass er selbst weiterführende Literaturhinweise sowie Hinweise zur Einordnung eines Wirtschaftswissenschaftlers und einer Rechtsphilosophin selber recherchieren muss. So richtig es ist, vom Konzept her einen Beitrag zu Ethik und Suchmaschinen in das vorliegende Handbuch aufzunehmen, so sehr muss die Vermittelbarkeit des Inhalts dieses Beitrages für den philosophiefremden Leser für eine künftige Auflage hinterfragt werden.

Mit den jüngst bekanntgegebenen Ethikstandards, an die sich Google, Microsoft und Yahoo zu halten gedenken, wäre ein praxisnahes Thema gegeben, dass einen philosophisch-ethischen Kommentar aus Expertensicht verdienen würde<sup>18</sup>.

#### Fazit

Die erste Auflage dieses Handbuchs ist ein guter erster Wurf. Man wünscht dem Buch den nötigen kommerziellen Erfolg, damit eine zweite überarbeitete Auflage ins Auge gefasst werden kann. Die Aufgabe, Wissenschaftler, Praktiker und Information Professionals gleichermaßen glücklich machen zu wollen, kommt der Quadratur des Kreises gleich. Hier kann eine noch eindeutigeren Ausrichtung des Handbuchs auf eines der drei Zielpublika Klarheit schaffen. Nicht alle Kapitel sind auf demselben hohen Niveau geschrieben, das der Herausgeber in seinen eigenen Beiträgen vorgibt. Sehr gut ist auch die Idee mit den Abstracts zu jedem Beitrag. Die Kapitel «Moral und Suchmaschinen» und «Enterprise Search» könnten durch eine striktere Fokussierung noch an Profil gewinnen. Ein Glossar würde den Handbuchcharakter verstärken und würde auch einen Beitrag zur Begriffsklärung leisten, die im Kapitel «Universal Search» vermisst wird.

Eine noch bessere Strukturierung (ähnlich wie etwa beim im August 2008 erschienenen *Buch Web-2.0-Dienste als Ergänzung zu algorithmischen Suchmaschinen*, in dem Dirk Lewandowski als

Mitherausgeber auftritt) würde den Handbuchcharakter für den Praktiker noch verstärken. Die Lektüre des Buches erweitert nicht nur den Horizont des Lesers in Bezug auf nicht so geläufige Suchmaschinen, sondern fördert auch das Problembewusstsein für den Umgang mit Suchmaschinen. Beruhigend ist die nach der Lektüre gewonnene Erkenntnis: Es gibt nicht nur Google. Die noch weitaus bessere Nachricht für den interessierten Leser ist, dass es nun ein Suchmaschinenhandbuch in deutscher Sprache gibt.

Kontakt: 35stevie61@bluewin.ch

search/Martin White Impressum: London: Facet Publishing, 2007; oder *Advances in information retrieval: 30th European Conference on IR Research: proceedings/ECIR 2008*, Glasgow, UK, March 30 – April 3, 2008; Craig Macdonald ... [et al.] (eds.) Impressum: Berlin: Springer.

11 Lewandowski, S. 228.

12 Op. cit., S. 231, Im Sinne von *scannen* [zu engl. to scan = abtasten, absuchen]. Duden – Wörterbuch der Szenesprachen, hgg. vom Trendbüro. Mannheim, Leipzig, Wien, Zürich: Dudenverlag 2001.

13 Duden, *Das Herkunftswörterbuch*, Mannheim, 2006.

14 Lewandowski, S. 316.

15 Op. cit., S.317.

16 Spiegel Heft 44/2008, S.76 f, und *Süddeutsche Zeitung* <http://www.sueddeutsche.de/computer/artikel/27/117896/>

17 Op. cit., S.321. Es darf spekuliert werden, ob der Autor in seinem Beitrag auf den Beitrag «Freedom of Choice» der *European Economic Review* aus dem Jahre 1988 anspielt oder ob hier auf das gemeinsame Werk von Sen und Nussbaum «The Quality of Life» (1993 Oxford University Press) angespielt wird.

18 <http://www.sueddeutsche.de/computer/13/315899/text/>

# Pérenniser le document numérique

Frédéric Sardet

Comme le rappelle l'introduction à cet ouvrage<sup>1</sup>, il n'est plus possible de considérer le cycle de vie des documents numériques sans aborder la problématique de leur conservation et la «pérennisation de son intelligibilité».

Agir donc vite et bien, par un ordonnancement adéquat de l'information serait en la matière un bon point de départ, comme le souligne Catherine Dhérent, qui rappelle les étapes obligées d'un référencement adéquat des documents numériques. Une occasion aussi pour elle de rappeler le poids croissant des *records managers* (y compris dans des services d'archives historiques) pour définir des règles de production et gérer les référentiels.

Pour celles et ceux qui méconnaîtraient les procédures de normalisation de l'archivage, il est conseillé de lire la contribution de Catherine Lupovici sur l'OAIS (Open Archival Information

System, ISO 14721). Une bonne introduction pour tous ceux qui s'interrogent sur la manière de mettre en œuvre un archivage numérique de long terme dans une acception très puissante, puisqu'il s'agit moins d'une durée que d'un rapport au changement, reproductible de manière infinie, mais par là même qui pose des exigences fortes en terme de veille.

Une autre perspective du livre concerne le rattrapage en quelque sorte de productions documentaires sinon sauvages du moins sans description ni organisation, un cas plutôt fréquent dans les organisations. Yannick Toussein en profite pour montrer des solutions imaginables par l'analyse de contenus des textes, l'extraction d'information ou l'annotation. Dans une autre contribution, Bertrand Couasnon développe la question de la numérisation rétrospective appliquée à des manuscrits, ce qui permet de présenter les espoirs mis dans l'automatisation de l'annotation des images et de sources anciennes comme l'état civil.

Autre objet analysé à l'occasion de ce séminaire, tout ce qui touche le web. Produire des documents numériques présente un coût. L'économie de la gestion numérique est analysée par le Ca-

nadien Jean-Michel Salaün, qui évoque les problèmes de l'édition et les transformations induites par la généralisation du Web 2.0. Julien Masanès, dans un monde virtuel où chacun peut publier dans une grande instabilité et une profusion d'instanciations ne garantissant pas pour autant la pérennité, présente les méthodes expérimentées depuis les années 1990 pour l'archivage du web (collecte et organisation). Problématique parallèle, la pérennisation des publications scientifiques dans un contexte d'accès libre (archives ouvertes).

Enfin, une intéressante présentation, par le juriste Philippe Bazin, de la signature électronique et du cadre légal français – condition supposée de l'authentification des documents numériques – dans laquelle il rappelle les normes pour le maintien d'une archive probante plutôt que «légale».

Un nouvel opus qui fait une fois encore honneur au travail mené en France notamment à l'INRIA pour prendre à bras le corps la problématique du numérique dans le cadre de l'information documentaire.

Contact: [frédéric.sardet@lausanne.ch](mailto:frédéric.sardet@lausanne.ch)

1 *Pérenniser le document numérique: séminaire INRIA 2-6 octobre 2006 – Ambroise, ouvrage coord. par Lisette Calderan, Bernard Hidoine et Jacques Millet, Paris: ADBS, 2006, 206 p.*