

NEUERE ENTWICKLUNGEN IM INFORMATIONSD- UND PUBLIKATIONSDWESEN
UND IHRE MÖGLICHEN AUSWIRKUNGEN

Ein Beitrag von Helge Steenweg

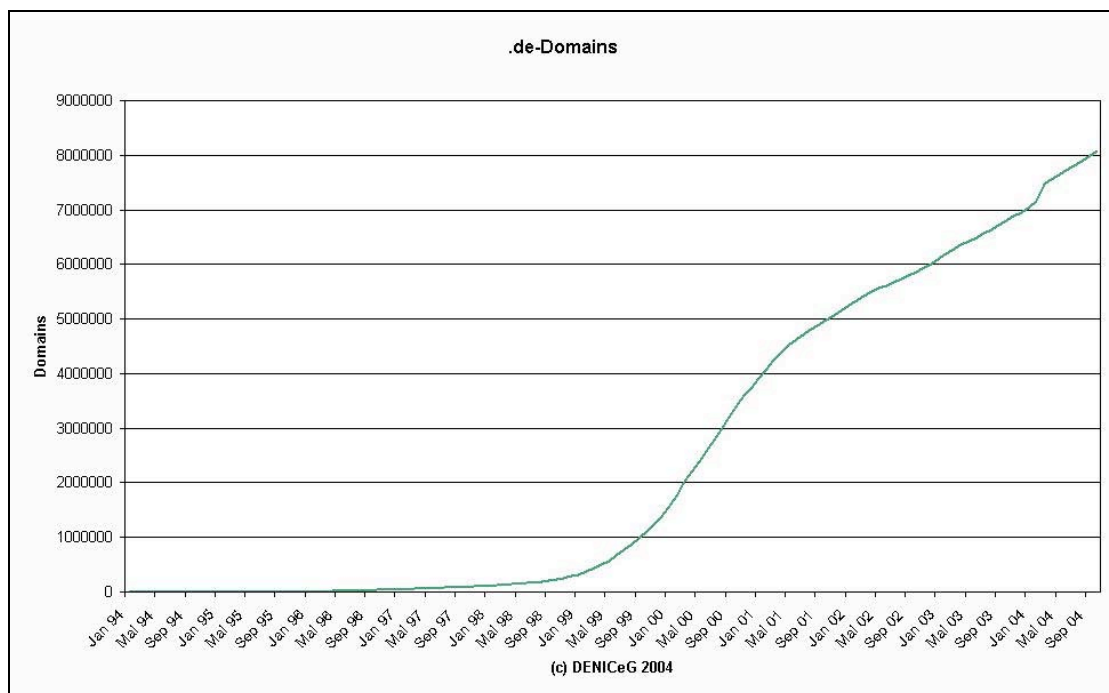


AUSGANGSLAGE: VERMITTLUNG VON WISSEN UND INFORMATIONEN

Die Verbreitung und der Bedarf an Informationen nimmt in den letzten Jahrzehnten überproportional zu. Dies ist natürlich zunächst eine subjektive Wahrnehmung, allerdings mag dies auch der Verfasser des Zedler-Universallexikon-Artikels „Weltveränderung“ zu seiner Zeit (1748) gedacht haben, als er notierte:

"Denn alle 40 Jahr, welche nach der Sündfluth für ein Mannes=Alter gerechnet werden, wird insgemein eine andere Welt, an Lebens=Art, Gemüths=Beschaffenheit, Principiis und Moden der Menschen. Wenn man volends vier mahl 40 Jahr, oder vier Generationen, das ist, 160 Jahr, zusammen nimmt, wird die Veränderung so groß, daß es auch ein Blinder mercken mögte."¹

Ohne die Vermittlung von Wissen und die Verbreitung von Informationen qualitativ bewerten zu wollen, zeigt doch ein Blick auf das Wachstum allein der deutschen Domains, in welchem rasantem Tempo sich der Wandel im Informationsangebot allein in dem letzten Jahrzehnt vollzogen hat (Grafik: DENIC, Stand 3.11.2004)²:



¹ Zitiert nach der online verfügbaren Ausgabe 1748 des Münchener Digitalisierungszentrums ([http://mdz.bib-bvb.de:80/digbib/lexika/zedler/images/ze54/@ebt-link;cs=default;ts=default;pt=2399180;lang=de?collection=images;book=ze54;target=IDMATC H\(entityref,ze540940\)](http://mdz.bib-bvb.de:80/digbib/lexika/zedler/images/ze54/@ebt-link;cs=default;ts=default;pt=2399180;lang=de?collection=images;book=ze54;target=IDMATC H(entityref,ze540940))), (man erreicht diese Seite auch etwas einfacher über die manuelle alphabetische Navigation zum Begriff "Weltveränderung").

² Vgl. <http://www.denic.de/de/domains/statistiken/domainentwicklung/index.html>.

Die rasche Zunahme des Informationsangebotes hat Folgen, die hier nur kurz angerissen werden können:

- Zunahme an Quantität bedeutet nicht zwangsläufig auch eine Zunahme an Qualität (Problematik der Quellenkritik, Problematik Spam)
- Zunahme an Quantität bedeutet nicht zwangsläufig zahlreichere und aussagekräftigere Ergebnisse bei der Suche (Problematik der effektiven Recherche)
- Lernorte und Lernmedien verändern sich (früher oftmals im Umfeld von Bildungsinstitutionen, jetzt verstärkt selbstorganisiertes Lernen am eigenen Computer mittels Internetzugang)
- Wissen verliert schneller an Gültigkeit (lebenslanges Lernen)
- allmähliches Ansteigen elektronischer Volltexte

DIE BETEILIGTEN AM PRODUKTIONSPROZESS

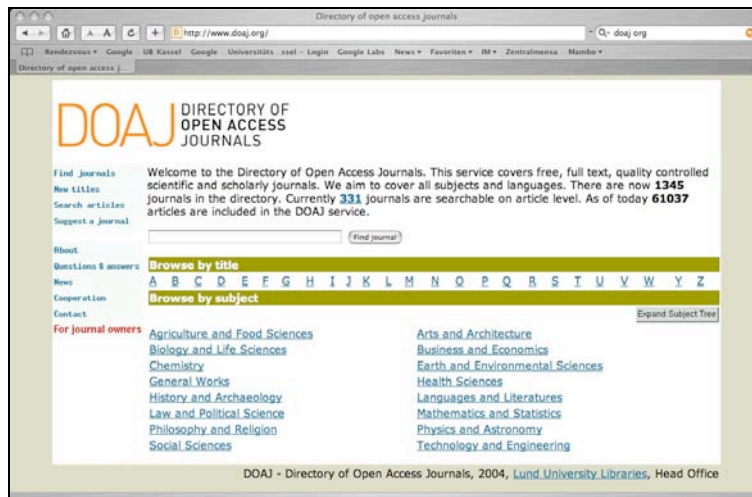
Nach herkömmlichem Produktionsprozess galt häufig die Abfolge: Autor-Verlag-Leser. Bedingt durch veränderte Produktionsbedingungen (z.B. Copyright-Problematik, Kostensteigerungen etc.) treten mittlerweile mehr oder minder deutliche Verschiebungen oder gar Veränderungen auf, die gerade in letzter Zeit bedingt durch Willensbekundungen im Umfeld der Open Access Initiative an Bedeutung gewinnen.

AUTOR

Fachwissenschaftler, verstärkt in den Naturwissenschaften, stehen unter dem Druck des "*publish or perish*" – sie müssen als Maßstab für wissenschaftliches Arbeiten und berufliches Fortkommen in renommierten, viel zitierten Zeitschriften publizieren. Da in den letzten Jahren immer offensichtlicher wird, dass sich das Publikationswesen nach herkömmlicher Machart und zu bestehenden Konditionen mit etablierten Wissenschaftsverlagen immer weniger finanzieren lässt, wird nach neuen Wegen gesucht.

So wächst die Anzahl der elektronisch verfügbaren Zeitschriften stetig. Das "*Directory of Open Access*" der Universität Lund³ verzeichnet derzeit 1345 Zeitschriften.

³ Vgl. <http://www.doaj.org>.



Ein anderes Beispiel ist die Public Library of Science (PLoS), die im Jahr 2000 von dem Nobelpreisträger Harold Varmus und anderen Prominenten gegründet wurde. PLoS verfolgt ein anderes Geschäftsmodell bei der Herausgabe von wissenschaftlichen Zeitschriften im Rahmen des freien Zugangs zu wissenschaftlichen Informationen, Open Access (OA). Im Jahre 2003 erschien die erste Ausgabe von PLoS Biology, einem online kostenfrei erhältlichen peer-reviewed Journal, das thematisch die gesamte Fachrichtung Biologie umfassen soll. Der kostenfreie Zugang wird dadurch ermöglicht, dass vom jeweiligen Autor ca. 1500 US\$ als Veröffentlichungsgebühr bzw. zur Deckung der Kosten für Redaktion und Begutachtung für das Erscheinen seines Aufsatzes erhoben werden⁴. Eine gedruckte Version der Zeitschrift kann für eine Jahresgebühr erworben werden.

Entscheidend für den weiteren Erfolg dieses Geschäftsmodells dürfte die Akzeptanz in Wissenschaftlerkreisen sein, die sich daran misst, inwieweit die Zeitschrift Aufnahme in die Fachreferenzierungen und Fachdatenbanken findet. Im Oktober diesen Jahres erschien erstmals PLoS Medicine, begleitet von einem gewissen Medienspektakel, das den Schluss zulässt, dass diese neue Publikationsform auch bei den etablierten Wissenschaftsverlagen ernster genommen wird.

⁴ Einige Institutionen unterstützen Autoren, wenn sie sich für eine Publikation in PLoS-Journals entscheiden, vgl. <http://www.plos.org/journals/model.html>.



Bedingt durch die allmählich einsetzende öffentliche Diskussion zu Open Access – vielleicht auch schlichtweg aus marketingstrategischen Überlegungen heraus – haben sich im Laufe des Jahres 2004 Verlagsaktivitäten entwickelt, die neue Finanzierungswege aufzeigen⁵. So hat Springer das „Open-Choice-Modell“⁶ seit Juli 2004 für Autoren im Angebot. Falls neben der regulären Druckpublikation eine weitere parallele Online-Ausgabe frei verfügbar sein soll, muss der Autor 3000 US\$ für verlegerische Dienste zahlen⁷.

Der Verlagskonzern Reed-Elsevier stellt seit Mitte des Jahres 2004 öffentlich Änderungen bei der Freigabe von Autorenrechten in Aussicht. Allerdings soll der Autor erst nach einer Frist von 6 Monaten seine Publikation entweder auf eigenen Webseiten oder aber über ein „institutional repository“⁸ zugänglich machen dürfen.

*"An author may post his version of the final paper on his personal web site and on his institution's web site (including its institutional repository). Each posting should include the article's citation and a link to the journal's home page (or the article's DOI). The author does not need our permission to do this, but any other posting (e.g. to a repository elsewhere) would require our permission."*⁹

Neben der Akzeptanz der Beteiligten dürfte ein entscheidender Fakt für das Gelingen von Open-Access-Geschäftsmodellen sein, inwieweit die Finanzierung solcher Veröffentlichungen unterstützt werden kann. Sicherlich lässt sich wie in einigen

⁵ Vgl. hierzu R. Poynder, No Gain without Pain, in: Information Today 21, Nr. 10, 2004, S.1-2.

⁶ Vgl. <http://www.springeronline.com/sgw/cda/frontpage/0,11855,1-40359-0-0-0,00.html>.

⁷ Vgl. <http://www.researchinformation.info/special2004springer.html>.

⁸ S. unten.

⁹ Aus einer E-Mail von Karen Hunter (Senior Vice President, Strategy, Elsevier, zitiert nach Henk Ellerman, Eintrag vom 1. Juni 2004, im Weblog „-In Between-“, vgl.: <http://eepi.ubib.eur.nl/iliit>.

Länder bereits angedacht auch an eine gesetzliche Regelung denken, jedoch wäre es sehr viel einfacher, wenn bei der Wissenschaftsförderung zukünftig ein entsprechender Etatposten vorgesehen bzw. in der Bewilligung von Projekten gefordert und verankert wird. Nur so ließe sich auch die derzeitige Skepsis von führenden Wissenschaftsverlagsvertretern entscheidend ändern.

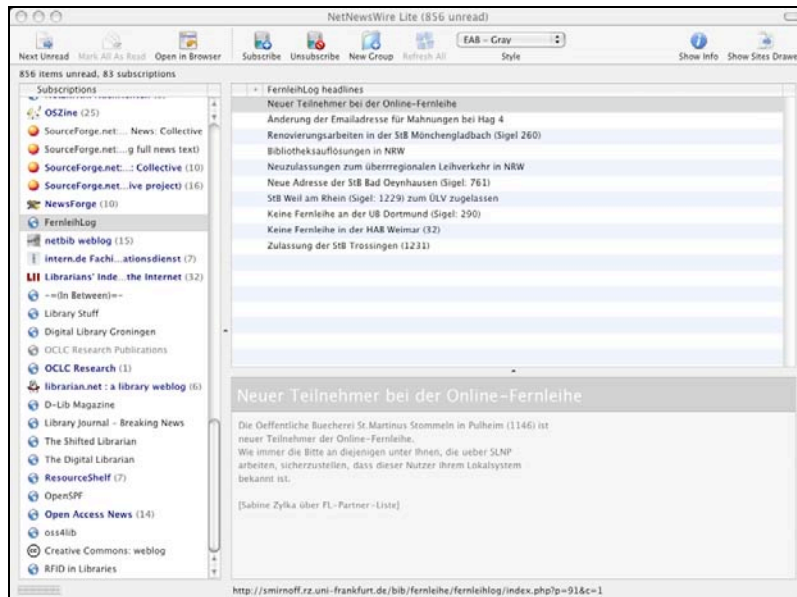
Eine andere Art ohne große technische Kenntnisse selbst zu publizieren und Informationen zu verbreiten, stellen die sogenannten Weblogs dar. Es handelt sich dabei um Webseiten, die von einer Person oder einem Personenkreis gepflegt und aktualisiert werden. Ursprünglich dienten Weblogs, die es seit etwa 1997 gibt, der einfachen Verbreitung persönlicher Inhalte (Gedanken, Meinungen), werden aber zunehmend themenorientiert zu wichtigen Informationsquellen im Web¹⁰. Da die Betreiber der Weblogs, Blogger genannt, oftmals zu den aktiveren Internetnutzern gehören, werden auf ihren Seiten sehr schnell Informationen verbreitet. Auch im Bibliothekswesen hat diese Informationsvermittlung seit einiger Zeit Einzug gehalten:



Um über Änderungen auf einer Webseite informiert zu werden, kann man die jeweilige Webseite als Bookmark abspeichern und regelmäßig besuchen. Es gibt jedoch eine weitere komfortablere Möglichkeit, die gerade auch im Zusammenhang mit Weblogs gerne genutzt wird: RSS¹¹, das ein plattformübergreifendes XML-basiertes Format darstellt; ursprünglich 1999 von der Firma Netscape entwickelt. Nutzer können mit einem entsprechenden RSS-Reader entsprechende Angebote abfragen und so sehr schnell Informationsangebote vorsortieren:

¹⁰ Vgl. <http://www.blogger.com/start>.

¹¹ Aufgelöst zu *Rich Site Summary* oder *Really Simple Syndication* oder *Rich Site Services*. Eine alternative Feed-Option stellt „Atom“, forciert von [blogger.com](http://www.blogger.com), dar.



Unter www.feedster.com existiert eine Suchmaschine für Weblogs und RSS-Feeds. Eine weitere neuartige Form des Publizierens stellt die WikiPedia (in Deutschland: <http://de.wikipedia.org/wiki/Hauptseite>) dar, eine Online-Form der Enzyklopädie, die in der Gemeinschaft ohne übergeordnete Kontrollinstanz oder Redaktion hergestellt und gepflegt wird. Die Wiki-Software ermöglicht gemeinsames Schreiben und Korrigieren von Texten.



Mittlerweile ist die kostenfrei recherchierbare WikiPedia akzeptiert und hat mit 100.000 Besuchern pro Tag deutlich mehr Zuspruch als kommerzielle ehemals etablierte Produkte. Seit Mitte Oktober 2004 ist ein CD-Image zum Brennen kostenfrei verfügbar, so dass die Enzyklopädie auch offline benutzbar ist. Gegen einen geringen Selbstkostenpreis kann auch eine CD bestellt werden.

Vermutlich werden die zukünftigen Folgen dieser neuen Publikationsformen (Weblogs, RSS, Wiki) derzeit noch etwas unterschätzt, jedoch ist nicht

auszuschließen, dass es in Verbindung mit der Open-Access-Bewegung und der flächendeckenden Versorgung mit institutionalisierten Dokumentenservern (*“institutional repositories“*), wie sie im Folgenden noch erläutert werden, zu einem Print- resp. Publikations-Paradigma kommen kann.

VERLAG

Einen Einblick in das deutsche Verlagswesen bietet das nunmehr zum dritten Mal vorliegende Branchenbarometer Elektronisches Publizieren¹² des Börsenvereins 2004, das jeweils zum Oktober des Jahres erscheint. Danach haben elektronische Produkte weiterhin hohes Umsatzpotential. Gegenüber den bisherigen Jahren hat sich die Erlössituation der befragten Verlage deutlich gebessert. Knapp 70% der Verlage gaben die Relation zwischen Kosten und Gewinnen als ausgeglichen an bzw. machten Gewinne. Auch die positive Einschätzung bezüglich des Online-Publishings hat sich gegenüber dem Jahr 2003 von 9% auf 18% nahezu verdoppelt¹³. Die durchgängige Erwartung fast aller Verlage geht von einem steigenden Anteil des Electronic Publishing (EP) mit ebenso steigenden Gewinnen in den nächsten Jahren aus. Bei den Verlageinschätzungen bezüglich der bisherigen Erwartungen und der zukünftigen Bedeutung einzelner Produkte ist sehr deutlich der hohe Stellenwert der CD-Rom heute erkennbar; zukünftig erscheint die CD jedoch nicht mehr, sondern wird von der DVD abgelöst. Absoluter Zukunftsträger scheint für die Verlage das Online-Publishing zu werden, das in der heutigen Einschätzung einen eher mittleren Erwartungswert hat. Gleichwohl heute als auch zukünftig wird der Internet-Buchhandel als zweitwichtigster Aspekt angesehen. Die Bedeutung des E-Learnings, das noch in der Befragung des Jahres 2003 zukünftig eine große Rolle spielen sollte, wird nunmehr nur noch mittelmäßig beurteilt. Die Einschätzung der Verlage bezüglich E-Books hat sich zwischen 2003 und 2004 nicht geändert; in beiden Untersuchungen rangieren E-Books ganz unten in der Erwartung. Der Umsatzanteil von elektronischen Produkten wird noch vorsichtiger als in den vorangehenden Jahren beurteilt.

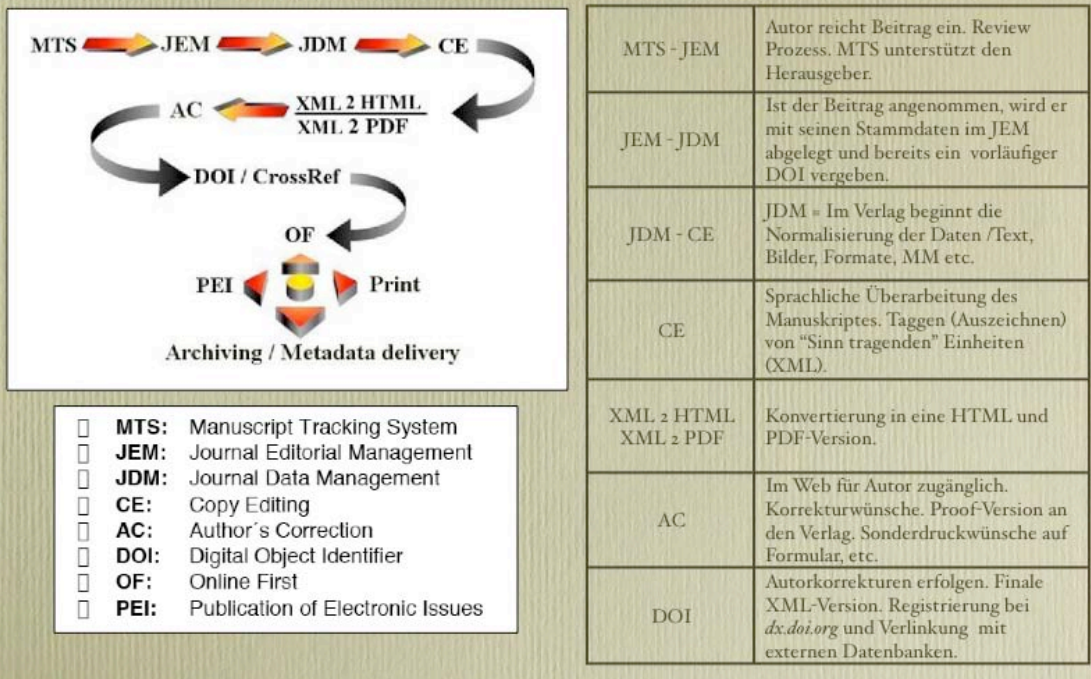
Bei der Beurteilung von digitalen Workflows hingegen ist ein deutlicher Anstieg der positiven Beurteilungen zu verzeichnen. Welche Arbeitsschritte bei der Herstellung eines Zeitschriftenartikels beispielsweise bei Springer vollzogen werden müssen, hat V. Böing in einem Vortrag im Rahmen der DGI-Online-Tagung „Competence in Content“ im Juni 2003 verdeutlicht¹⁴:

¹² Branchenbarometer 2004 des Börsenvereins, vgl. unter <http://www.boersenverein.de/sixcms/detail.php?id=69181&rubrik=69174>.

¹³ s.o. Branchenbarometer 2004, S. 3.

¹⁴ Erreichbar über: <http://www.boersenverein.de/de/69181?id=69181&rubrik=69176&skip=25>.

Herstellung eines Zeitschriftenartikels (Springer)



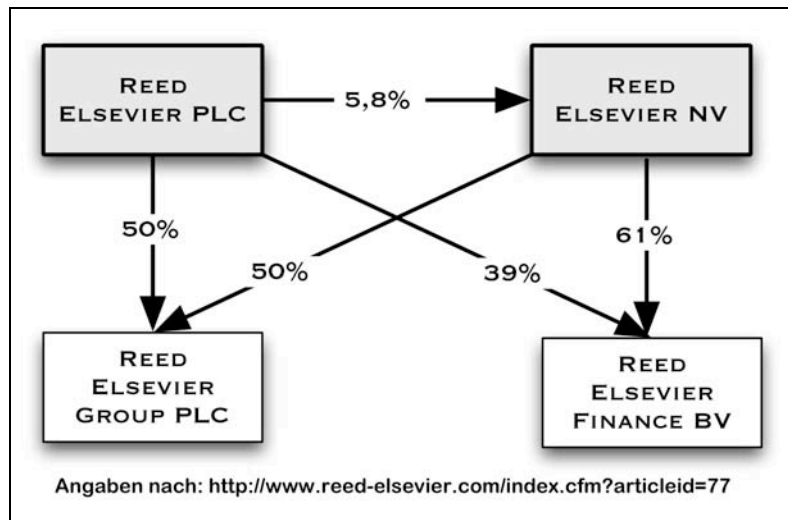
FINANZEN UND MONOPOLISIERUNGEN

Sehr deutlich geht aus dem Springer-Workflow die Verzahnung von Print- und Digital-Publikation hervor. Der gesamte Produktionsprozess ist unter finanztechnischen Gesichtspunkten optimiert. Dies spricht auch aus den Presseverlautbarungen, wenn Springer beispielsweise im Juli 2004 Finanzgebern Gewinnmargen von bis zu 20 Prozent verspricht und sich auf die *"lukrativen Bereiche Wissenschaft, Technik und Medizin"* beschränken will, da hier *"traumhafte Renditen von bis zu 70%"* möglich sind¹⁵.

Ebenso deutlich wird die Ausrichtung auch bei dem Konzern Reed-Elsevier, einem der „Global Player“ im Informationswesen. Auf den eigenen WWW-Seiten kann man einen Einblick bekommen, wie Reed-Elsevier mittlerweile strukturiert wurde¹⁶:

¹⁵ Vgl. Yahoo-Finzen vom 2. Juli 2004, unter: <http://de.biz.yahoo.com/040702/299/43m2n.html>.

¹⁶ Vgl. <http://www.reed-elsevier.com/index.cfm?articleid=77>.



Reed-Elsevier Group plc besitzt demnach alle Geschäfte im Publikations- und Informationsbereich, Reed Elsevier Finance nv kümmert sich um die finanziellen Belange. Nach den internationalen Fusionen und Käufen¹⁷, wie Ideal (Harcourt und Academic Press), Urban&Fischer, Spektrum Akademischer Verlag, Beilstein Vertriebsgesellschaft und letztlich Lexis-Nexis als führender Online Anbieter für Wirtschafts- Presse- und Rechtsinformationen steht der Konzern nach eigener Aussage auf vier Säulen:



In einem Finanzbericht¹⁸ anlässlich der Benelux Conference der Deutschen Bank vom 15. September 2004 sieht sich die Firma gemessen an der Gesamtentwicklung

¹⁷ Seit 1985 als Geschäftsstrategie „Globale Expansion“.

¹⁸ Das Folgende nach einem Vortrag von Marc Armour auf der Benelux Conference der Deutschen Bank am 15. September 2004 : http://www.reed-elsevier.com/media/pdf/p/2/Deutsche_Bank.pdf.

in den jeweiligen Geschäftsbereichen gut positioniert¹⁹:

Jährliches Wachstum	Markt	Reed-Elsevier
Elsevier	3	6
LexisNexis	2	3
Harcourt	5	9
Reed Business	-10	-3

In dem für Bibliotheken und sonstige Informationsanbieter wichtigen Geschäftsfeld von Reed Business steht der o.a. durchschnittliche Rückgang im Wachstum dennoch nicht mit einem Verlust in Zusammenhang, wie die Gewinnerwartung der Jahre 1999 bis 2004 ausweist:



Neben der ausgewiesenen prozentualen Steigerungsrate sollte man sich klarmachen, dass die Einnahmen in fünf Jahren auf über das 8fache der Einnahmen des Jahres 1999 gestiegen sind.

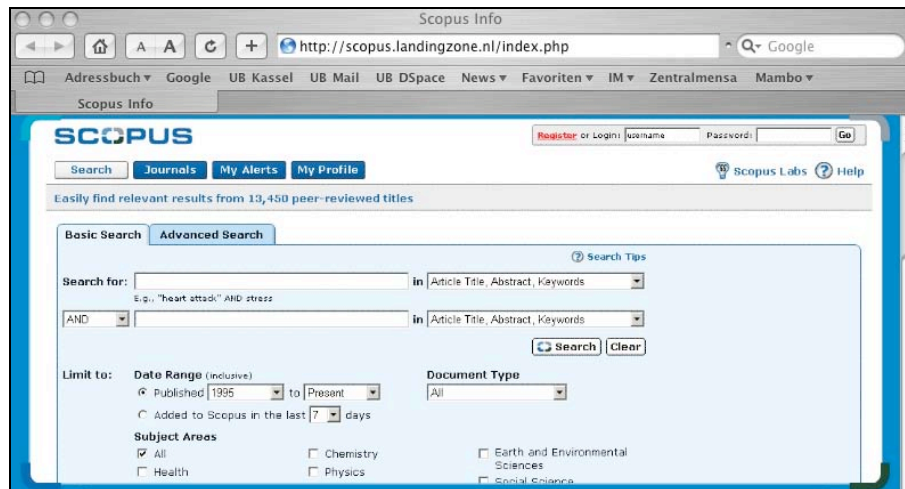
Reed-Elsevier kann man sicherlich als sogenanntes „Data-Warehouse“ ansprechen. Es werden die Produkte, die erzeugt werden, gleichfalls über eigene Portale (Scirus, ScienceDirect) vermarktet. Ab dem 3. November 2004 wurde von Reed-Elsevier ein weiteres Produkt, SCOPUS²⁰, eine Abstracts- und Index-Datenbank, die mit den jeweiligen Volltext über CrossRef²¹ verlinkt ist, auf den Markt gebracht. Die Textsuche soll intuitiv und einfach sein. Derzeit hat der Nutzer Zugriff auf 14.000 peer-reviewed Journale von mehr als 4000 Verlegern, einige davon sollen bis auf

¹⁹ Vgl. http://www.reed-elsevier.com/media/pdf/p/2/Deutsche_Bank.pdf.

²⁰ Derzeit unter: <http://www.info.scopus.com>.

²¹ Vgl. <http://www.crossref.org>.

Jahrgänge aus den 60er Jahren zurückgehen. Eingebunden sind weiterhin 400 Open-Access-Journale.



Reed-Elsevier hat durch seine fortschreitende Globalisierung eine relative Marktführerschaft erreicht und hat durch die Nutzung neuer Vertriebswege es geschafft, Inhalte umfassend weltweit und zeitnah anzubieten. Dies mag für die jeweiligen Finanziers und Aktienbesitzer sehr vorteilhaft sein, die Kunden werden die Annehmlichkeiten, die ein solches monopolartiges Angebot ermöglicht, letztlich auch bezahlen müssen. Hier Änderungen zu erwarten wäre unrealistisch. Änderungen können sich nur durch eine andere Art der Informationsvermittlung und –Produktion ergeben. Dies dürfte auch den jeweiligen Verlagskonzernen klar sein und erklärt die deutlichen Reaktionen auf die einsetzende Open-Access-Bewegung.

ANBIETEN VON DIGITALEN BÜCHERN

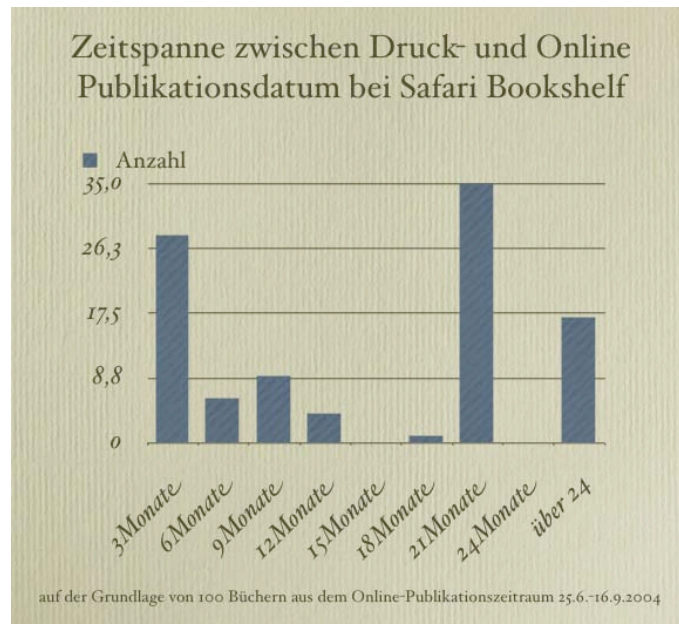
Bücher selbst anzubieten und digital zu vermarkten ist sicherlich keine neue Geschäftsidee, gewinnt jedoch, wenn man sich an die oben angeführten Ausführungen der Börsenvereinsbefragung bezüglich der zukünftigen Stellung des Internet-Buchhandels erinnert, deutlich auch für Verlage an Bedeutung und Interesse. So hat O'Reilly bezogen auf die Zielgruppe der DV-Programmierer und IT-Profis eine elektronische Referenz-Bibliothek namens „Safari Bookshelf“ auf Subskriptionsbasis aufgebaut. Dieses Unternehmen ist ein Joint-Venture zwischen O'Reilly and Associates und der Pearson Technology Group (mit Addison-Wesley, Cisco Publishing, New Riders, Peachpit Press, Prentice Hall, Que und Sams).



Derzeit beinhaltet diese digitale Buchhandlung („*online bookstore*“) mehr als 2000 Bücher, die nach Kategorien oder Volltextsuche durchsucht und kapitelweise gelesen werden können. Es soll sich ausschließlich um neuere Publikationen (seit 2000) handeln. Zur Nutzung wird eine monatlich zu zahlende Subskription benötigt, die sich nach der Größe des eigenen „Bücherbords“ (*bookshelf*) bemisst (Größe: Safari Max/Safari Basic oder Starter-Paket). Man wählt interessante Bücher aus und stellt sie online auf sein Bücherbord, bis es virtuell gefüllt ist. Möchte man mehr Bücher dort stehen haben, erwirbt man die nächst höhere Subskriptionsstufe und kann dann weiter auffüllen. Eine andere Möglichkeit besteht darin, Bücher abzuwählen und anschließend durch neue zu ersetzen. Allerdings muss ein abgewähltes Buch mindestens 30 Tage auf dem Online-Bücherbord gestanden haben. Hat man bei einer Suche interessante Titel gefunden, kann man bereits an der Darstellung der Ergebnisse erkennen, welche Möglichkeiten man als Leser hat. Sofern das Buch auf dem eigenen Bücherbord steht, kann man das komplette Buch ansehen, anderenfalls erhält man nur eine kurze Voransicht von einigen Zeilen. Erst durch das Hinzufügen dieses Buches zum Bücherbord wird es wieder gänzlich lesbar. Weiterhin sind kapitelweise Downloads erlaubt, die sich zahlenmäßig an der monatlichen Subskription orientieren. Das Safari-Max-Paket kostet beispielsweise 19,99 US\$.

Untersucht man die dort angebotenen Bücher nach der Zeitspanne zwischen Erscheinungsdatum der jeweiligen Print- und Online-Ausgabe, so ergibt sich ein recht heterogenes Bild²²:

²² Ausgewertet wurden 100 Bücher des Online-Publikationszeitraumes 25. Juni – 16. September 2004.



Deutlich sind die Spitzen bei zeitnahen (3 Monate Differenz) und zeitfernen (21 Monate) Zeitspannen erkennbar. Vermutlich lässt sich dieses Ergebnis nur auf die unterschiedlichen Verlage und Verlagspolitiken zurückführen, denn eine Aufschlüsselung nach den Verlagen ergibt folgendes Bild:

Differenz in Tagen	Mittelwert	Median
O'Reilly	104,9	44,5
Premier Press	663,2	596,0
Que	228,2	128,0
Muska&Lipman	574,5	596,0

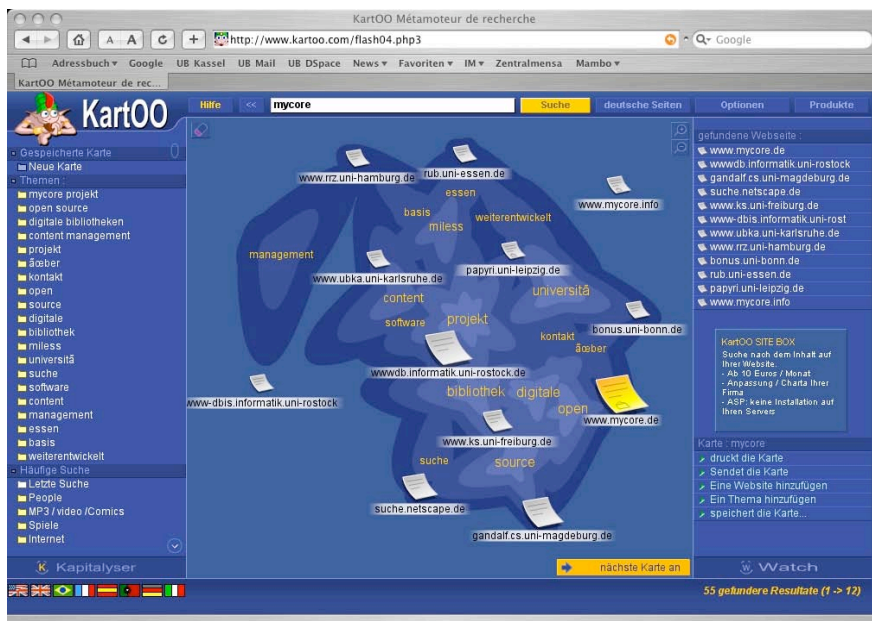
Offenkundig ist O'Reilly als ausführender Verlag des Safari Bookshelves stark daran interessiert, zeitnah Publikationen in dieses Abonnement-Verfahren einzubringen, um es attraktiv zu gestalten. Da bei einigen Verlagen einige Ausreißer das Bild des Mittelwertes etwas verfälscht hätten, wurde zusätzlich der Median errechnet. O'Reilly hat einige Bücher gleichzeitig mit der Druckausgabe online bereitgestellt.

INFORMATIONSMITTLER

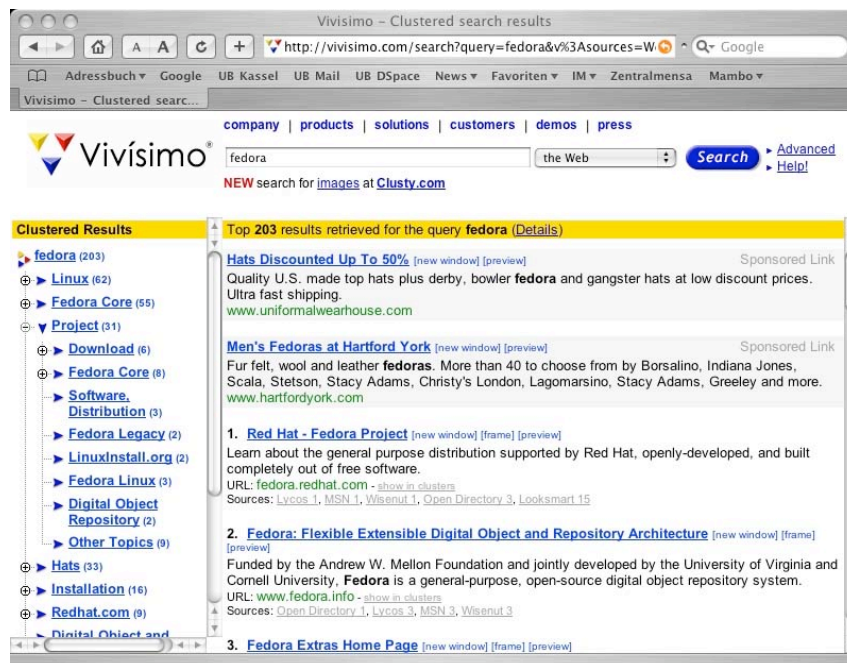
NEUE TECHNIKEN

Um der Fülle an manchmal auch unerwünschten Informationen Herr zu werden, gibt es immer wieder Versuche, die Ergebnisse von Google oder Yahoo zu filtern oder zu strukturieren. Zwei nicht ganz so bekannte Meta-Suchmaschinen versuchen, Ergebnisse von anderen Suchmaschinen zusammenzufassen und dem Nutzer zur Verfügung zu stellen. Sie erledigen diese Aufgabe allerdings in der Darstellung sehr unterschiedlich. Kartoo (www.kartoo.com) präsentiert die Suchergebnisse visuell,

indem es Beziehungen zwischen den Suchbegriffen darstellt.



Eine weitere Meta-Suchmaschine mit Cluster-Technologie stellt Vivisimo (www.vivisimo.com) dar. Vivisimo fasst bei einer Suche ähnliche Begriffe zusammen und ermöglicht so einen schnellen Überblick über die Suchergebnisse. Dabei dürfte von Vorteil sein, dass man gerade bei mehrdeutigen Antworten die Varianten schnell erkennen und auswählen kann ohne langwierig Suchergebnisse Seite für Seite abzurufen.



Im Beispiel ist im linken Bereich schnell erkennbar, dass der Suchbegriff „Fedora“ nicht nur eine Linux-Distribution abdeckt sondern gleichzeitig auch für ein

Dokumentenserver-Projekt benutzt wird. Das Aufklappen des Begriffsbaumes ermöglicht eine schnelle Navigation zu dem gewünschten Projekt.

KOOPERATIONEN MIT BIBLIOTHEKEN UND BIBLIOTHEKSVERBÜNDEN

Betrachtet man die Suchgewohnheiten von Bibliotheksnutzern in OPAC's, so kann man sich eigentlich nicht mehr über den Erfolg von Suchmaschinen wundern. In Kassel ergab eine statistische Untersuchung²³ der OPAC-Abfragen über das Jahr 2002, dass fast 90% der Abfragen nur maximal drei Suchschlüssel (Titel, Schlagwort, Person) überhaupt nutzen. Immerhin 97% der Nutzer begnügen sich mit der Angabe eines Suchbegriffes. Daher ist nicht verwunderlich, dass die Suchmaschinenanbieter bei der Erschließung weiterer Betätigungsfelder auch den Bibliothekssektor einbeziehen wollen. So stellt OCLC²⁴ in einem Pilotprojekt namens „Open WorldCat“²⁵ zwei Millionen Datensätze aus der insgesamt 55 Millionen umfassenden WorldCat-Datenbank den Suchmaschinen von Google und Yahoo zur Verfügung. Bei den extrahierten Daten soll es sich um die am meisten frequentierten Daten des OCLC-Datenbestandes handeln. Bei Google sind die Daten seit Dezember 2003, bei Yahoo seit Mai 2004 recherchierbar.

Die Nutzung dieses Dienstes erfolgt, indem man neben dem Suchbegriff zusätzlich die Suche wie folgt einschränkt:

site:worldcatlibraries.org Suchbegriff

worldcatlibraries Suchbegriff

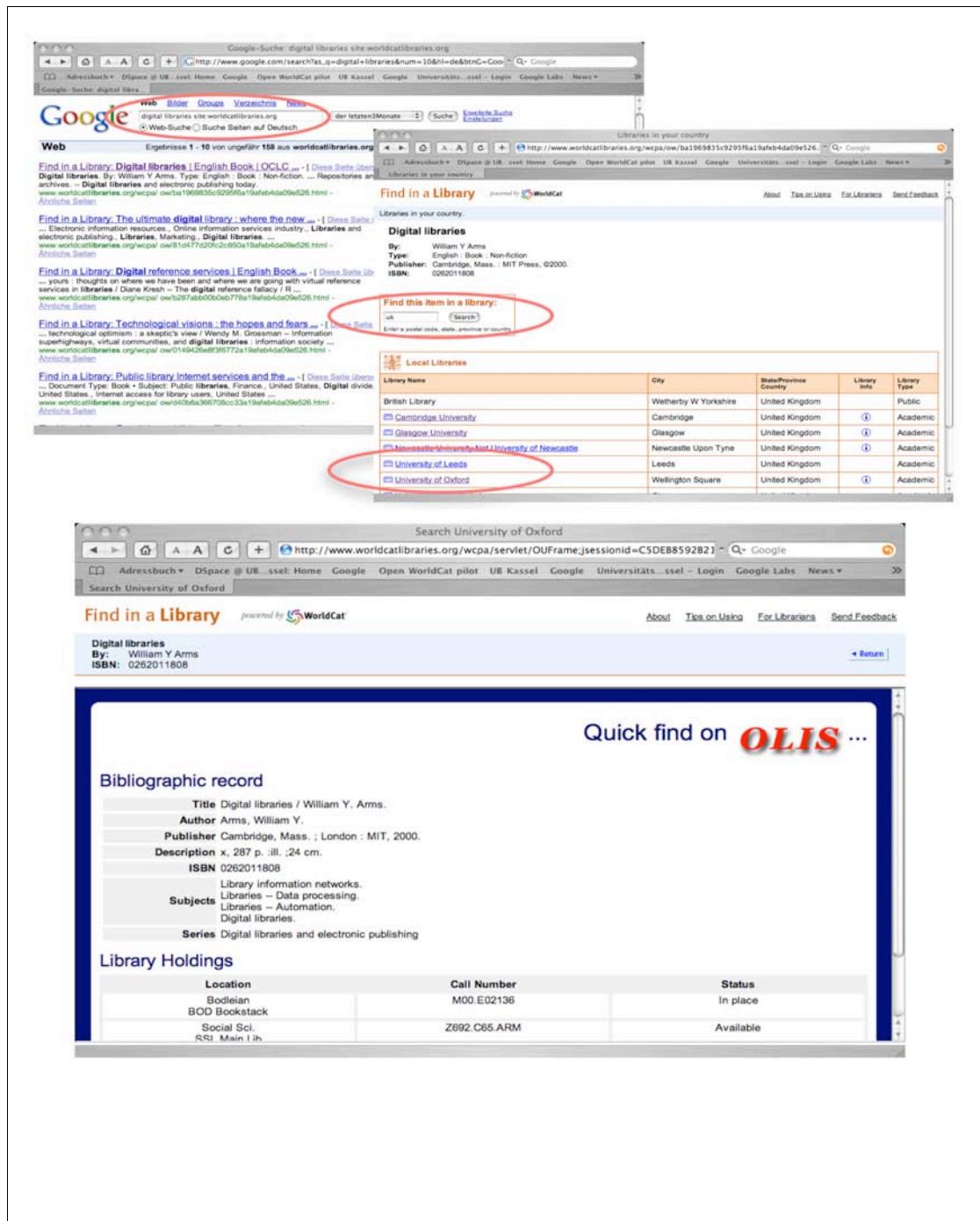
In der Google-Anzeige werden die Fundstellen gekennzeichnet mit: „Find in a library“. Beim Anklicken einer Fundstelle wird eine Link-Seite geöffnet, die weitere regionale Einschränkungen zulässt²⁶. Ein weiterer Klick verbindet den Suchenden mit dem entsprechenden OPAC.

²³ Ich danke Helmut Gick, Kassel, für diese Informationen.

²⁴ Vgl. www.oclc.org.

²⁵ Vgl. <http://www.oclc.org/worldcat/open/default.htm>.

²⁶ Vgl. N. O'Neill, Open WorldCat Pilot: A User's Perspective, in Searcher, The Magazine for Database Professionals 12, Nr. 10, 2004, S. 52-61.

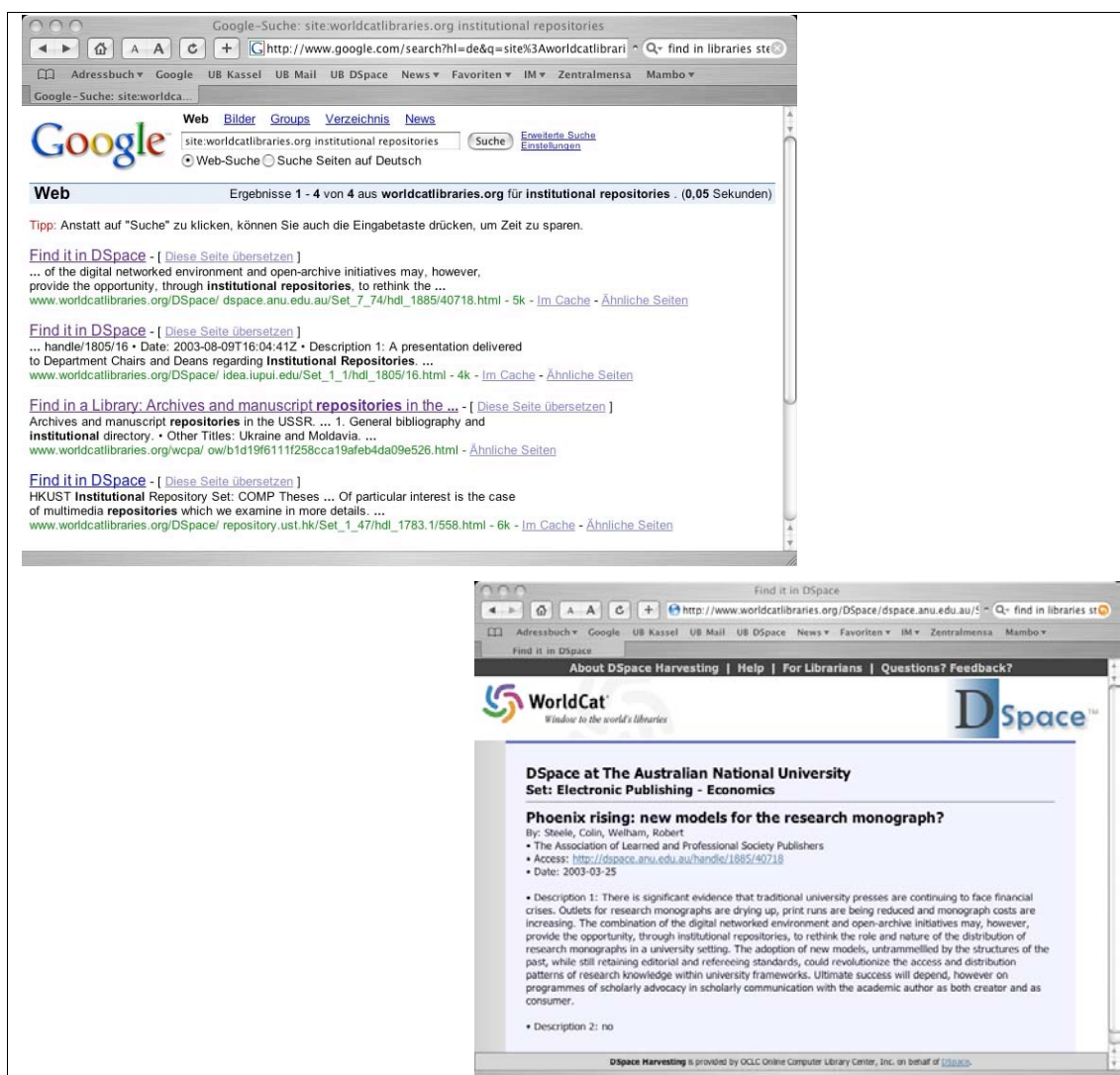


Die Nutzung scheint den OCLC-Angaben zufolge Erfolg versprechend. Zwischen Januar und September 2004 zählte OCLC 3,14 Millionen Links auf den OpenWorldCat²⁷. Immerhin 8% der Nutzer ließen sich von der Überblicksseite anschließend mit einem OPAC verbinden.

Bleibt nur abzuwarten, wie sich diese derzeit noch frei zugängliche Dienstleistung nach der Pilotphase finanziell und zugangstechnisch entwickelt. Für Bibliotheken dürfte diese Entwicklung beobachtungswürdig sein, da sie sowohl Chancen als auch Risiken birgt.

²⁷ Vgl. http://www.oclc.org/worldcat/open/open_worldcat_factsheet.pdf

Nahe liegend in diesem Zusammenhang ist zudem für Suchmaschinenbetreiber die Erschließung von Volltextarchiven. Im April 2004 wurde eine Kooperation von 125 Universitäten und Google zur Suche mittels Google in DSpace-Archiven abgeschlossen²⁸. Es ist geplant, eine entsprechende Option in der Google eigenen „Erweiterten Suche“ einzubauen. Ergebnisse erhält man bereits jetzt, wenn man entsprechend der oben geschilderten Vorgehensweise im OpenWorldCat sucht. Der einleitende Text lautet dann entsprechend: „Find it in DSpace.“



KOOPERATION MIT VERLAGEN UND AUTOREN

Google-Print

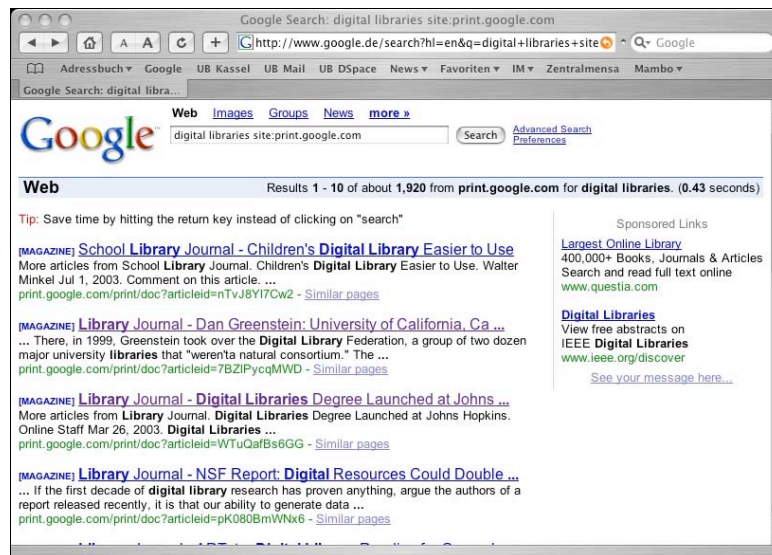
Seit Oktober 2004 bietet Google einen weiteren Service: „Google Print“²⁹. Verleger

²⁸ Vgl. <http://google.blogspot.com/archives/001203.html>. DSpace ist eine sog. Institutional Repository-Software, entwickelt von HP und dem MIT – vgl. <http://www.dspace.org>.

²⁹ Vgl. <http://print.google.com/googleprint/about.html>.

oder Autoren können ihre Publikationen über Google gescannt und indiziert erhalten. Bei der Suche soll dem jeweiligen Nutzer immer nur ein kleiner Teil der Informationen zur Verfügung gestellt werden; das Copyright bleibt beim Rechteinhaber. Ein entsprechendes Angebot hat Google unter der Adresse „<https://print.google.com/publisher/>“ gemacht. Einige Zeitschriftenaufsätze und Abstracts (scheinbar aus der angekündigten Kooperation mit Reed-Elsevier) sind bereits jetzt recherchierbar, indem man die Suche einschränkt:

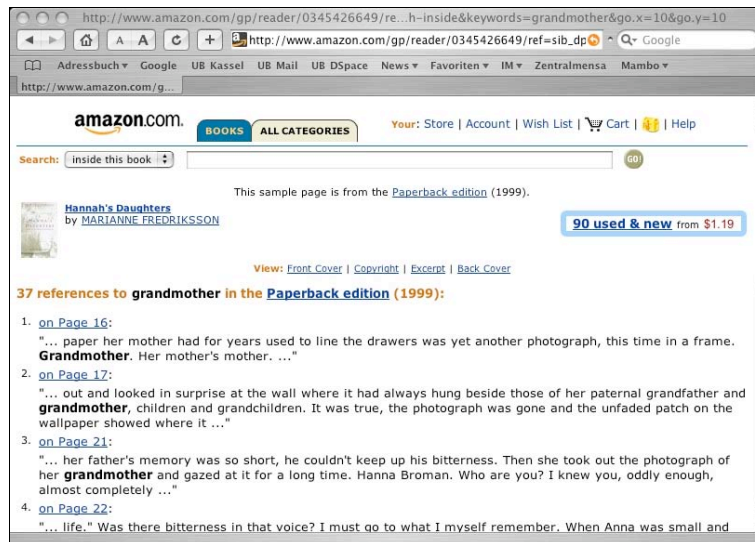
site:print.google.com Suchbegriff



NEUE ANBIETER

A9

Amazon weitet seine Dienstleistungen beständig aus. Bereits seit geraumer Zeit ist es möglich, zu einigen Büchern, die man über die Suchfunktion gefunden hat, Inhaltsverzeichnisse und Textproben vorab anzusehen. Weiterhin ist oftmals bei diesen Büchern eine Suche in den Textinhalten möglich:



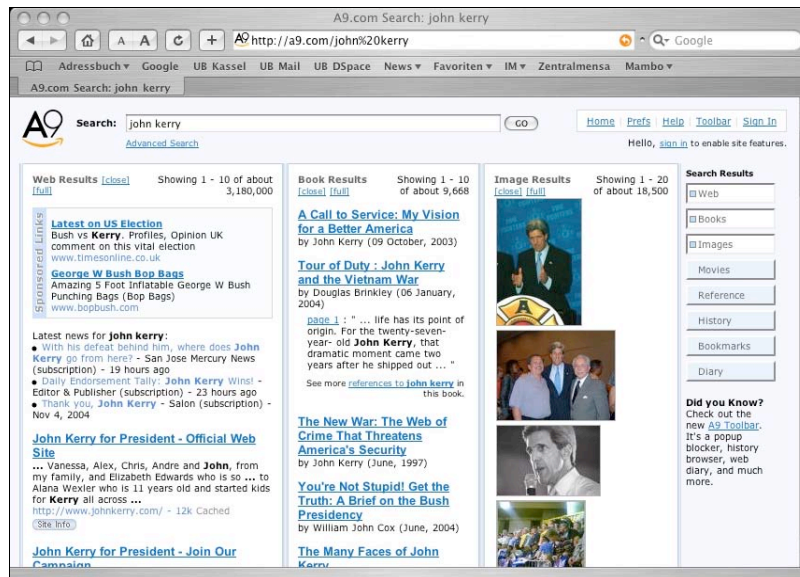
Eine zumindest für Pica-Bibliotheken und -Bibliotheksnutzer nützliche Funktion bietet John Udell mit seinem Bookmarklet-Service³⁰ an. Über diese Seite ist ein Bookmarklet, das in fast allen Browsern lauffähig ist, zu erstellen, das es ermöglicht, im Amazon-Katalog gefundene Bücher automatisch ohne jegliche Eintragungen über die ISBN-Suche im eigenen PICA-OPAC als Bestand zu überprüfen³¹.

Seit Mitte September verfügt Amazon zudem über eine eigene Suchmaschine (a9.com). Amazon greift für die Internet- und Bild-Suche auf Google, bei der Buch-Suche aus nahe liegenden Gründen allerdings auf eigene Ressourcen zurück. Alle Ergebnisse werden in konfigurierbarer Spaltenform dargestellt. Als Service wird neben der Suchfunktion eingetragenen Benutzern ermöglicht, Bookmarks abzuspeichern und zu editieren, Links zu speichern, Notizen zu Seiten zu machen, etc³².

³⁰ Vgl. <http://weblog.infoworld.com/udell/stories/2002/12/11/librarylookupGenerator.html>.

³¹ Das bei Udell angegebene Beispiel der TU Hamburg-Harburg funktioniert bei den meisten anderen PICA-OPAC's entsprechend. Bei einigen nicht standard-konfigurierten Installationen ist manuelle Nacharbeit nötig, so z.B. für die UB Kassel: `javascript:var%20re=/([V-]|is[bs]n=)(\d{7,9}[\dX])/i;if(re.test(location.href)==true){var%20isbn=RegExp.$2;void(win=window.open('http://opac.bibliothek.uni-kassel.de/DB=23/CMD?ACT=SRCHA&IKT=8520&SRT=YOP&TRM='+isbn,'LibraryLookup','s crollbars=1,resizable=1,location=1,width=575,height=500'))}`

³² Zu den einzelnen Features vgl. <http://a9.com/-/company/whatsCool.jsp>.



OPEN ACCESS

Im Zusammenhang mit neuen Geschäftsmodellen bei der Publikation von Beiträgen in Fachzeitschriften wurde bereits auf die Open-Access-Bewegung hingewiesen, die es sich zur Aufgabe gemacht hat, freien Online-Zugang zu wissenschaftlicher Literatur zu schaffen³³. In Deutschland hat sich der Diskussionsprozeß seit der „Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen“ vom 22. Oktober 2003 intensiviert³⁴.

Wichtig wäre jetzt, dass zum einen auf breiter Front ein Mentalitätswandel bei den Autoren wissenschaftlicher Publikationen einsetzt, der die bisherigen bestehenden Akzeptanzprobleme von OA-Journalen beseitigt. Zum anderen muss die Infrastruktur geschaffen werden, damit diese Online-Dokumente auch entsprechend angeboten, gefunden und archiviert werden, also schlicht hin für Produzenten auch akzeptabel erscheinen. Bleibt es zudem bei den derzeitigen rechtlichen Überlegungen und den Ankündigungen einiger bedeutender Verlagsgruppen (z.B. Reed-Elsevier), dem Autor zumindest nach einem Zeitraum von einem oder einem halben Jahr zu gestatten, seine Arbeit entweder auf einen eigenen Server oder einen zentralen Dokumentenserver („*institutional repository*“) (IR) zu legen, müssen diese Dokumentenserver vorhanden sein und sollten technisch in der Lage sein, diese Informationen optimal zu erschliessen (Volltextindizierung, übergeordnete Meta-Suchmaschinen) und langfristig zu sichern (Langzeit-Archivierungsproblematik). Diese Problematik besteht weltweit und hat bereits zu zahlreichen Diskussionen und praktischen Umsetzungen geführt. In Deutschland existieren ebenfalls Projekte dieser Art³⁵. Institutional Repositories werden zukünftig eine große Funktion für das Veröffentlichen von wissenschaftlicher Literatur haben³⁶.

³³ Vgl. <http://www.soros.org/openaccess/g/read.shtml>.

³⁴ Vgl. http://www.mpg.de/pdf/openaccess/BerlinDeclaration_dt.pdf.

³⁵ Vgl. <http://www.mycore.de>, <http://urmel.uni-jena.de>.

³⁶ Vgl. M.A. Drake, Institutional Repositories: Hidden Treasures, in Searchers, Vol. 12, Nr. 5 - May 2004 – vgl. <http://www.infotoday.com/searcher/may04/drake.shtml>.

OPENREP-INITIATIVE

Stellt man einige Kriterien auf, so erkennt man sehr rasch, dass es eine umfassende Lösung derzeit nicht gibt³⁷. Bevor jedoch ein neues zeitintensives Projekt aufgesetzt wird, um eine wie auch immer umfassende Neuentwicklung voranzutreiben, wäre es möglicherweise sinnvoller, aus vorhandener Open-Source-Software ein modular aufgebautes „Open Repository“ (OpenRep) aufzubauen, das es Institutionen ermöglicht, entweder eine lokalisierte Variante komplett zu übernehmen oder aber nur Teile modular an vielleicht schon vorhandene eigene Dokumentenserver über Standardschnittstellen (z.B. OAI-PMH) anzugliedern. So wäre es denkbar, dass Module, wie z.B. die oftmals fehlende Volltextindizierung³⁸, entwickelt werden, die an die unterschiedlichen heute schon existenten IR-Varianten angeschlossen werden können und so jede Institution eine auf sich zugeschnittene Lösung erhalten kann. So könnte vergleichsweise schnell ein Netzwerk von Institutional Repositories entstehen, deren Inhalte mittels vernetzter Suchmaschinen schnell gefunden werden können und ebenso selbstverständlich auch längerfristig erreichbar sind.

Schwierig aber nicht minder wichtig ist die Frage der langfristigen Archivierung der Dateninhalte dieser Dokumentenserver. Bislang haben nur wenige Institutional Repositories entsprechende Funktionen. Hier besteht in absehbarer Zeit massiver Nachholbedarf. Eine von OCLC und RLG ins Leben gerufene Arbeitsgruppe (PREMIS) hat im September 2004 einen ausführlichen Bericht zu dieser Problematik vorgelegt: *Implementing Preservation Repositories for Digital Materials: Current Practice and Emerging Trends in the Cultural Heritage Community*³⁹.

FAZIT

Neuere Untersuchungen bestätigen die Annahme, dass die Nutzung des Internets als Informationsquelle immer wichtiger wird⁴⁰. Auch langfristig angelegte Studien zeigen deutlich die Ausweitung des Informationscharakters. So scheinen Ergebnisse auf der Basis von anonymisierten Suchanfragen aus sechs Jahren einen deutlichen prozentualen Rückgang der Sex-orientierten Abfragen von 20% im Jahre 1997 auf heute 5% in den USA bzw. 8-10% in Europa zu belegen bei einem gleichzeitigen Anstieg von Anfragen zum E-Commerce (Informationen über neue Autos, Bücher etc.)⁴¹. Die Analyse der Abfragegewohnheiten bzw. –Techniken hat sich in den der Studie zugrunde liegenden sechs Jahren nur unwesentlich verändert: Im Durchschnitt wurde in den Suchmaschinen mit maximal zwei Worten pro Anfrage und maximal zwei Anfragen pro Session gesucht. Bei der Betrachtung der Suchergebnisse war in der Regel nur die erste Seite entscheidend.

Diese Vorgehensweise, die in abgewandelter Form auch bei OPAC-Abfragen

³⁷ Einen sehr guten Überblick bietet das nunmehr in dritter Auflage vorliegende „A Guide to Institutional Repository Software, August 2004, unter: <http://www.soros.org/openaccess///software>.

³⁸ Nur wenige IR-Software verfügt über eine Volltextindizierung, meist mittels Lucene - vgl. auch <http://jakarta.apache.org/lucene/docs/index.html>.

³⁹ Vgl. <http://www.oclc.org/research/projects/pmwg/surveyreport.pdf>.

⁴⁰ Vgl. z.B. zuletzt The Digital Future Report: Surveying the Digital Future, Oktober 2004 unter: <http://www.digitalcenter.org/downloads/DigitalFutureReport-Year4-2004.pdf>.

⁴¹ A. Spink, B. Jansen, Public Searching of the Web (Information Science & Knowledge Management, Vol. 6), 2004.

nachzuweisen ist⁴², weist aber auf ein großes Problem hin. Der Glaube, mit einer Abfrage alle in Frage kommenden Antworten erhalten zu haben, muss relativiert werden. Suchmaschinen vom Typ Google sind zwar ein hervorragendes Suchhilfsmittel, jedoch gibt es wohlmöglich daneben noch andere Informationen. Auch die Vorstellung, dass aus politischen oder finanziellen Gründen irgendwann und irgendwo Suchergebnisse manipuliert werden können, ist nicht völlig weltfremd.

Suchmaschinen haben im Internet eine zentrale Schlüsselfunktion, sie werden – wie gezeigt – immer weiter ausgebaut und mit neuen Funktionen versehen. Gleichzeitig ist zu bedenken, dass dieser Ausbau finanziert werden muss und die Finanzierung durch Investoren Abhängigkeiten schaffen kann. Um der Gefahr einer starken Kommerzialisierung und möglichen Risiken beim Ranking zu entgehen, sollte der Ansatz der Schaffung von miteinander vernetzten regionalen Open-Source-Suchmaschinen zumindest im akademischen Bereich forciert werden⁴³.

Die Studie des Digitalcenter weist weiter darauf hin, dass über 50% der befragten Internetnutzer die Inhalte im Web für ganz oder fast immer richtig und vertrauenswürdig halten⁴⁴. Aufgeschlüsselt nach spezifischen Anbietern gelten offizielle Verwaltungs- oder Regierungsseiten zu 73%, Seiten etablierter Firmen sogar zu 74,4% als vertrauenswürdig, anders als bei der Wertung privater Webseiten, den nur zu 9,5% vertraut wird⁴⁵. Der kritische Umgang mit Rechercheergebnissen und Webseiten dürfte zukünftig ein weiterer Themenschwerpunkt bei der Vermittlung von Informationskompetenz sein.

ANSCHRIFT DES AUTORS:

Dr. Helge Steenweg

Informationsmanagement UB Kassel, Diagonale 10, D-34127 Kassel

E-Mail: steenweg@bibliothek.uni-kassel.de

⁴² Vgl. die statistischen Ergebnisse aus der Kassler OPAC-Analyse oben.

⁴³ Siehe die Initiative zur Förderung der Suchmaschinentechologie und des freien Wissenszugangs unter: <http://www.suma-ev.de>.

⁴⁴ Vgl. <http://www.digitalcenter.org/downloads/DigitalFutureReport-Year4-2004.pdf>, S. 52.

⁴⁵ Vgl. <http://www.digitalcenter.org/downloads/DigitalFutureReport-Year4-2004.pdf>, S. 53.