

von Michael Mönnich

### Einführung

Das zehnjährige Jubiläum der Zeitschrift B.I.T.online im Jahre 2007 soll als Anlass dienen, einen Blick zurück auf die Entwicklung der Informationstechnik im Bibliotheksbereich zu werfen. Naturgemäß ist eine solche Rückschau stets eine subjektive Auswahl, in diesem Falle aus der Sicht des Universitätsbibliothekars an einer Technischen Hochschule. Neben der reinen Rückschau wird auch eine Extrapolation der Trends in die Zukunft gewagt.

Als Einstiegspunkt für die Zeitreise mag das Thesenpapier "Zukünftige Anforderungen an die Informationsinfrastruktur der Universität Karlsruhe 1996 - 2005" dienen, das folgende Eckpunkte definiert:

1. Für die zukünftige Forschung an der Universität Karlsruhe (TH) wird der Zugriff auf Informationen in elektronischer Form wesentlich, die Papierform wird an Bedeutung verlieren.
2. Die elektronische Information wird an den Arbeitsplatz des Wissenschaftlers geliefert werden.
3. Für eine Übergangszeit wird die Bibliothek die Nachfrage sowohl nach Print- als auch nach elektronischen Medien befriedigen müssen.
4. Die Bibliothek wird zudem ihre Dienstleistungen der sich wandelnden Situation anpassen. Sie wird Zugangspunkte für die über Netze zugängliche, für die Universität relevante Information schaffen (Filterfunktion) und die elektronischen Informationen sichten bzw. aufbereiten (Qualitätssicherungsfunktion).
5. Die Bibliothek wird zudem den langfristigen Zugriff auf relevante elektronische Informationen gewährleisten (Archivierung).
6. Bislang liegt noch kein standardisiertes Modell für die Bereitstellung von Information in elektronischer Form vor, das die Anforderungen der Wissenschaft nach Qualität, Zitierfähigkeit und sicherem Zugriff befriedigt.
7. Die Bibliothek sollte einen Dokumentenserver aufbauen, der die wissenschaftliche Produktion der Universität zur Verfügung stellt (Verlagsfunktion für Hochschulschriften).

Diese Thesen wurden 1995 vor dem Hintergrund der rasanten Ausbreitung des Internets und der damit verbundenen netzgestützten Informationen und Kommunikationsmöglichkeiten in den 1990er-Jahren formuliert. Wichtige Meilensteine waren zu Beginn der 1980er-Jahre die Einführung von TCP/IP als Kommunikationsprotokoll und 1989 der Entwurf von HTML. 1984 wurde an der Universität Karlsruhe die erste deutsche E-Mail empfangen, 1986 entstanden die ersten .de-Domains, darunter uka.de und dortmund.de. 1991 wurde das Informationssystem Gopher an der Universität von Minnesota entwickelt und im Europäischen Kernforschungslabor CERN das WWW realisiert. Ein Jahr später folgte der erste Webbrowser (Mosaic). 1993 gab es 500 Webserver weltweit, davon 15 in Deutschland, 1994 entstanden die ersten Suchmaschinen und die ersten deutschen Zeitschriften gingen

online. 1995-1997 zeigte der Browserkrieg zwischen Netscape und Microsoft, welchen Stellenwert das Internet inzwischen erreicht hatte.

Die Bibliothek hat diesen Wandel durch die Einführung neuer Dienste kontinuierlich nachvollzogen: 1993 nahm die Universitätsbibliothek Karlsruhe ihren Gopherserver in Betrieb, 1994 folgte der über Telnet weltweit zugängliche Olix-OPAC. 1995 löste den Gopherserver ein WWW-Server ab, über den auch im OPAC recherchiert werden konnte. 1996 wurde der KVK ins Leben gerufen, das Ausleihsystem in den Web-OPAC integriert und Benachrichtigungen an Benutzer als E-Mail versandt. Ein Jahr später startete das elektronische Volltextarchiv EVA, 1998 Subito und der lokale Scandienst LEA. Im Jahr 2000 nahm das Digitale Videoarchiv DIVA den Betrieb auf, 2003 wurde das Recommendersystem Bibtip in den OPAC integriert. Im folgenden Jahr gründete die Universitätsbibliothek den Universitätsverlag Karlsruhe auf Basis von Open Access und 2005 wurde auf Basis von Suchmaschinenteknik und Open Source Tools ein neuer Katalog entwickelt (XOPAC).

Die Liste zeigt, dass Bibliotheken die neuen Entwicklungen aufnahmen und - entgegen dem landläufigen Bild als einer etwas verstaubten Einrichtung - zeitnah umsetzten und den Herausforderungen des neuen Informationszeitalters mit der Einrichtung von Webservern, WebOPACs und E-Mail begegneten. Von den zahlreichen Veränderungen der letzten Jahre seien die folgenden besonders hervorgehoben:

1. Content ins Web - Aufbau von Volltextservern
2. Fachinformation - von der CD zum Webangebot
3. Der Aufstieg der elektronischen Zeitschriften
4. Open Access - ein goldener Weg für Bibliotheken?
5. Digitalisierung - Papier ins Netz
6. Erschließung im Internetzeitalter - Fehlanzeige?
7. Suchmaschinen - ordnen die Information im Netz, machen sie zugänglich und nutzbar

#### Content ins Web - Aufbau von Volltextservern

Der Aufbau von Volltextservern an Bibliotheken begann 1997 mit dem Ziel, die in der Universität produzierten Inhalte zentral nachzuweisen und zu archivieren. Waren es zunächst nur wenige Universitäten - darunter Dortmund, Karlsruhe, Leipzig und Stuttgart - so griffen die meisten Bibliotheken diese Idee auf und richteten eigene Archive ein. Der etwas altbackene Begriff "Volltextserver" wurde inzwischen durch "Institutional Repository" ersetzt. Kam es mancherorts noch zu Kompetenzstreitigkeiten mit anderen Einrichtungen der Universität wie den Rechenzentren, so hat sich inzwischen von wenigen Ausnahmen abgesehen die Bibliothek als Ort für Angebot und Archivierung von digitalen Volltexten etabliert. Der Deutschen Initiative für Netzwerkinformation (DINI) zufolge gibt es in Deutschland inzwischen 128 Dokumentenserver, die systematisch die Publikationen und andere Materialien ihrer Forscher frei zugänglich machen. Weltweit sind derzeit 1080 Repositorien im (inoffiziellen) Directory of Open Access Repositories (OpenDOAR) registriert. Deutschland liegt weltweit damit hinter den USA auf dem zweiten Platz vor England und Japan ein. Der Metakatalog OASE (<http://www.ubka.uni-karlsruhe.de/oase/>), der 1999 als Karlsruher Virtueller Volltextkatalog ins Leben gerufen wurde, verzeichnet heute 31 Volltextserver, die Bielefelder Suchmaschine BASE (<http://base.ub.uni-bielefeld.de/>) nennt 41 Quellen mit über 7,6 Millionen Dokumenten.

Deutsch

English

Budapest Open Access Initiative

Petition  
for guaranteed public access to public-funded research results

Suchbegriffe eingeben

Freitext

Titel  Körperschaft   
 Autor  Erscheinungsjahr

Volltitelanzeige:  Neues Fenster

Kataloge auswählen Auswahl

<p><b>Deutschland: Hochschulen</b></p> <ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Aachen, RWTH</li> <li><input checked="" type="checkbox"/> Augsburg, Uni</li> <li><input checked="" type="checkbox"/> Bamberg, Uni</li> <li><input checked="" type="checkbox"/> Bayreuth, Uni</li> <li><input checked="" type="checkbox"/> Berlin, TU</li> <li><input checked="" type="checkbox"/> Bielefeld, Uni</li> <li><input checked="" type="checkbox"/> Braunschweig, TU</li> <li><input checked="" type="checkbox"/> Chemnitz, Uni</li> <li><input checked="" type="checkbox"/> Cottbus, BTU</li> <li><input checked="" type="checkbox"/> Dortmund, Uni</li> <li><input checked="" type="checkbox"/> Erlangen-Nürnberg, Uni</li> <li><input checked="" type="checkbox"/> Freiburg, PH</li> <li><input checked="" type="checkbox"/> Freiburg, Uni</li> <li><input checked="" type="checkbox"/> Frankfurt, Uni</li> <li><input checked="" type="checkbox"/> Giessen, Uni</li> <li><input checked="" type="checkbox"/> Göttingen, Uni</li> <li><input checked="" type="checkbox"/> Hamburg, Uni</li> <li><input checked="" type="checkbox"/> Hamburg, H.-Schmidt-Uni.</li> </ul>	<ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Harburg, TU</li> <li><input checked="" type="checkbox"/> Heidelberg, Uni</li> <li><input checked="" type="checkbox"/> Hohenheim, Uni</li> <li><input checked="" type="checkbox"/> Kaiserslautern, Uni</li> <li><input checked="" type="checkbox"/> Karlsruhe, Uni</li> <li><input checked="" type="checkbox"/> Karlsruhe, FZK</li> <li><input checked="" type="checkbox"/> Köln, FH</li> <li><input checked="" type="checkbox"/> Konstanz, HTWG</li> <li><input checked="" type="checkbox"/> Konstanz, Uni</li> <li><input checked="" type="checkbox"/> Leipzig, Uni</li> <li><input checked="" type="checkbox"/> Mannheim, Uni</li> <li><input checked="" type="checkbox"/> Mainz, Uni</li> <li><input checked="" type="checkbox"/> Marburg, Uni</li> <li><input checked="" type="checkbox"/> Nürtingen-Geislingen, HWU</li> <li><input checked="" type="checkbox"/> Passau, Uni</li> <li><input checked="" type="checkbox"/> Potsdam, Uni</li> <li><input checked="" type="checkbox"/> Regensburg, Uni</li> <li><input checked="" type="checkbox"/> Reutlingen, HSB <b>NEU</b></li> <li><input checked="" type="checkbox"/> Saarbrücken, Uni</li> </ul>	<ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Siegen, Uni</li> <li><input checked="" type="checkbox"/> Stuttgart, HdM ePub</li> <li><input checked="" type="checkbox"/> Stuttgart, Uni</li> <li><input checked="" type="checkbox"/> Trier, Uni</li> <li><input checked="" type="checkbox"/> Tübingen, Uni</li> <li><input checked="" type="checkbox"/> Ulm, Uni <b>NEU</b></li> <li><input checked="" type="checkbox"/> Weimar, Bauhaus Uni</li> <li><input checked="" type="checkbox"/> Weingarten, PH</li> <li><input checked="" type="checkbox"/> Würzburg, Uni</li> </ul> <p><b>Deutschland: Sonstige</b></p> <ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Die Dt. Bibliothek</li> <li><input checked="" type="checkbox"/> GBV Online Ressourcen</li> <li><input checked="" type="checkbox"/> Publica der Fraunhofer-Gesellschaft</li> <li><input checked="" type="checkbox"/> Virtuelle Fachbibliothek Psychologie (Uni des Saarlandes)</li> </ul> <p><b>Ausländische Dokumentenserver</b></p> <ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> DiVA Portal Skandinavien</li> <li><input checked="" type="checkbox"/> MIT DSpace</li> <li><input checked="" type="checkbox"/> WU Wien: Dissertationen</li> </ul>
--	--	--

Suche nach der URN

Rückgabeoptionen für URNs:  Dokument  URL-Liste  
Was bedeutet URN? Der URN-Resolver Der Deutschen Bibliothek (mit URN-Beispielen)

OASE: <http://www.ubka.uni-karlsruhe.de/oase/>

Die Open Archives Initiative (OAI) legte zu Beginn des neuen Jahrtausends den Grundstein für die Vernetzung dieser Repositorien. Das OAI Protocol for Metadata Harvesting (OAI-PMH) wurde zur Übertragung von Metadaten eingerichtet und von einer wachsenden Anzahl von Institutionen unterstützt. Es bietet die Möglichkeit, neben den Metadaten auch die Volltexte selbst von den dezentralen Servern abzugeben und so Suchdienste wie OAIster der University of Michigan (<http://oaister.org/>) oder BASE aufzubauen.

Fachinformation - von der CD zum Webangebot

Der Umschwung von den bislang über CD-ROM verteilten Fachdatenbanken hin zu webbasierten Systemen begann 1997. Für die Bibliotheken bedeutete dies das Ende der mit zum Teil großem technischen Aufwand betriebenen CD-ROM-Server-Installationen (zum Beispiel von Holthaus und Heinisch) und eine Entlastung von deren personalaufwendiger

Pflege. Die traditionellen Fachinformationszentren wie DIMDI oder STN haben leider erst spät den Trend zu Internet und webbasierten, endnutzerorientierten Angeboten erkannt und hielten zu lange an ihren überkommenen Retrievalsprachen wie STN-Messenger u.ä. fest. Auch werden immer noch viele Fachdatenbanken produziert, deren WWW-Nutzerschnittstellen umständlich und unergonomisch sind.

Den Bibliotheken fällt dann die undankbare Aufgabe zu, diese Systeme ihren Google-verwöhnten Nutzern anzupreisen. Das Widerstreben der Studenten und Wissenschaftler, sich darauf einzulassen, wird dann häufig zum Anlass genommen, diesen mangelnde Informationskompetenz zu attestieren. Die Bibliotheken versuchen dem dann abzuhelpfen, indem sie Kurse zur Vermittlung von Informationskompetenz anbieten. Ob diese Anstrengungen Früchte tragen werden, scheint indes fraglich. Den Nutzern wäre es sicherlich am liebsten, würden die Fachinformationsanbieter ihre unbestreitbar wertvollen Daten in die den Nutzern vertrauten Suchmaschinen und Portale integrieren. Möglicherweise wäre eine Abkehr vom bisherigen Refinanzierungsmodell, bei dem sich die Fachdatenbankanbieter die Recherche in der Datenbank bezahlen lassen, der richtige Weg. Am Beispiel von PubMed zeigt sich, dass dadurch die Daten eine enorme Verbreitung finden. Alternativ zu einem solchen staatsfinanzierten wären auch Open Access Modelle denkbar, bei denen der Eintrag in die Datenbank bezahlt wird, wie beim örtlichen Telefonbuch.

#### Der Aufstieg der elektronischen Zeitschriften

Anders als die Datenbankhosts waren die großen Zeitschriftenverlage bei der Umstellung ihres Angebotes auf ein internetbasiertes Angebot erfolgreicher. Dies erstaunt, waren die Verlage doch viel mehr den Printmedien verhaftet als die Anbieter von Fachdatenbanken. Möglicherweise lag es wie so oft am Geld: Die erheblichen Mittel, die zum Aufbau eines umfangreichen und hochwertigen Webangebots an elektronischen Zeitschriften notwendig waren, wurden Ende der 1990er-Jahre durch starke Preiserhöhungen bei STM-Zeitschriften aus den Bibliotheksetats abgeschöpft. So bieten inzwischen fast alle großen Wissenschaftsverlage die Mehrzahl der Titel in elektronischer Form an. Die zur Anfangszeit des elektronischen Publizierens von manchen Kunden erhofften Kostenreduktionen bei Zeitschriftenpreisen sind indes ausgeblieben. Stattdessen waren die Bibliotheken gefordert, diese neuen Informationsquellen adäquat zu erschließen, den Nutzern nahe zu bringen und die diversen Lizenzmodelle zu bedienen. Eine große Hilfe war dabei die Einrichtung der EZB an der Universitätsbibliothek Regensburg (<http://ezb.uni-regensburg.de/>), die heute unentbehrlich ist für die Verwaltung von elektronischen Zeitschriftendaten. Bibliotheken organisierten sich in Konsortien auf Länder-, Bundes- oder Institutionenebene, um die meist in Paketen angebotenen elektronischen Journals günstig einzukaufen.

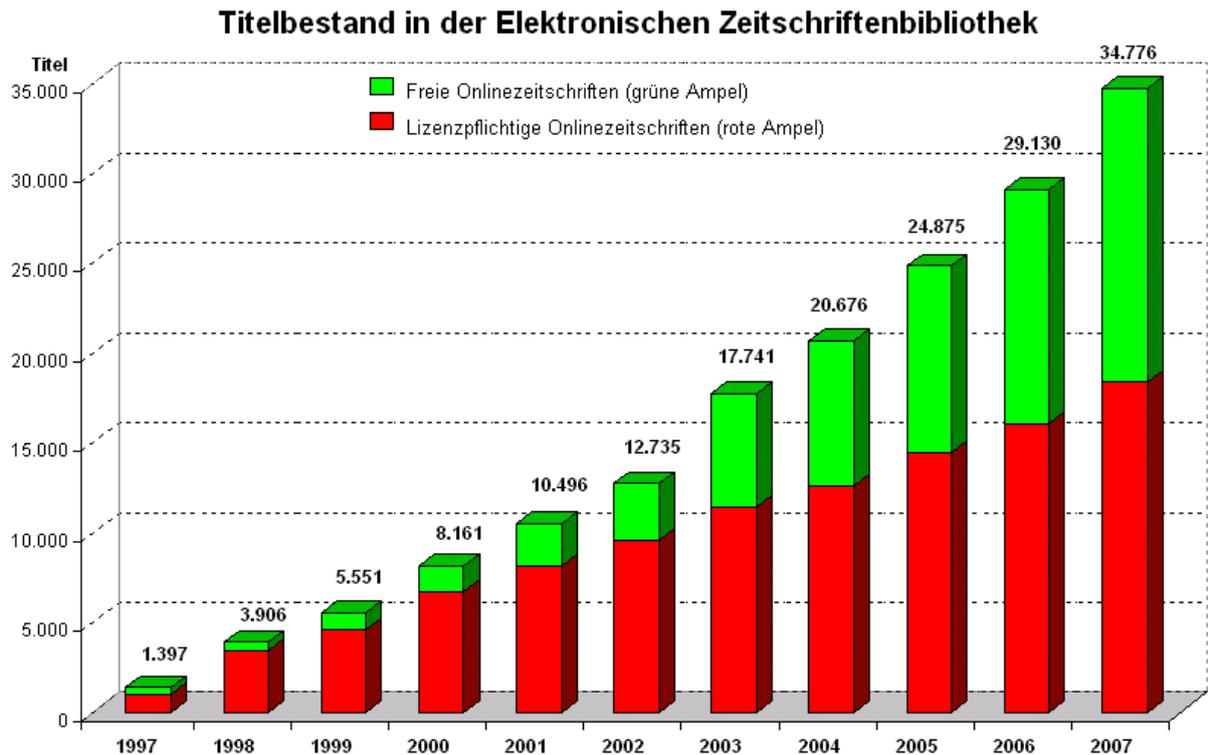


Abb. mit freundlicher Genehmigung der EZB

Ein weiterer Meilenstein war sicherlich die Aushandlung von Nationallizenzen für E-Journals durch die Deutsche Forschungsgemeinschaft im Jahr 2004. Dadurch erhielt der Trend hin zu elektronischen Zeitschriften einen nachhaltigen Schub. Insbesondere rückten damit die mit dem Umstieg auf e-only bei Zeitschriften verbundenen organisatorischen Vorteile verstärkt ins Bewusstsein der Bibliothekare. Man denkt seitdem verstärkt über die Aussonderung von den nun über Nationallizenzen abgesicherten alten Zeitschriftenbänden nach. Dem Aussondern älterer Titel steht allerdings noch die häufig schlechte Qualität der durch Retrodigitalisierung erzeugten Images von älteren Zeitschriftenheften entgegen.

#### Open Access - ein goldener Weg für Bibliotheken?

Parallel zu dem von den großen Wissenschaftsverlagen vorangetriebenen Umstieg auf elektronisches Publizieren nahm die Kritik an den Preissteigerungen zu. Vor allem von Seiten der Wissenschaftler und deren Gesellschaften sah man in den Preissteigerungen und daraus resultierenden Kündigungen von Zeitschriftenabonnements eine Bedrohung der Versorgung der Wissenschaftler mit der notwendigen Information und sah in der verlagsunabhängigen, elektronischen Publikation und der kostenlosen Verteilung über das Internet eine Alternative. Als "Urvater" dieser Idee kann sicherlich arXiv.org gelten, ein Server für Preprints aus den Bereichen Physik, Mathematik und Informatik, der 1991 startete und heute über 460.000 Dokumente enthält. NCSTRL (Networked Computer Science Technical Reports Library) war ein frühes Projekt der Universitäten Virginia University und Virginia Tech mit dem Ziel, technische Berichte aus den Fachgebieten Informatik und Elektrotechnik dezentral vorzuhalten und zu verteilen. NCSTRL begann 1994 und wurde bis 2001 betrieben.

Anfang des neuen Jahrtausends entstand dann die Open Access Initiative, als deren formaler Startpunkt die Budapest Open Access Initiative (<http://www.soros.org/openaccess>) gelten kann, die während einer Tagung in Budapest im Dezember 2001 gegründet wurde. Vorläufer

der Open Access-Bewegung in Deutschland waren die Bemühungen der IuK Initiative (Information und Kommunikation) der wissenschaftlichen Fachgesellschaften in Deutschland, deren Vertreter für alternative Modelle für das elektronische Publizieren im Wissenschaftsbereich eintraten. Für Deutschland besonders wirkmächtig war dann die Berliner Erklärung vom 23. Oktober 2003

([http://www.mpg.de/pdf/openaccess/BerlinDeclaration\\_dt.pdf](http://www.mpg.de/pdf/openaccess/BerlinDeclaration_dt.pdf)), die alle Forscher und Leiter von Forschungsprojekten aufrief, die Ergebnisse ihrer Arbeiten im Internet frei verfügbar zu machen und von allen großen Wissenschaftsorganisationen in Deutschland unterzeichnet wurde, namentlich von der Hochschulrektorenkonferenz, dem Wissenschaftsrat, der Deutschen Forschungsgemeinschaft, der Max-Planck-Gesellschaft, der Fraunhofer-Gesellschaft, der Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz, der Helmholtz-Gemeinschaft sowie vom Deutschen Bibliotheksverband.

Weitere Meilensteine waren die Open Access Stellung der Inhalte der weltgrößten Datenbank für medizinische Informationen als Pubmed (<http://www.ncbi.nlm.nih.gov/PubMed/>) im August 1997 durch die U.S. National Institutes of Health (NIH) und 2000 der Start des kommerziellen Open Access Verlages BioMed Central (<http://www.biomedcentral.com/>).

Von den Verlagen erwartungsgemäß von Anfang an heftig bekämpft, scheint sich Open Access als feste Größe auf dem Markt für wissenschaftliche Informationen zu etablieren. Zu nennen sind hier der Start von PubMed Central (<http://www.pubmedcentral.nih.gov/>) oder die Public Library of Science (PloS, <http://www.plos.org/>) in 2003. Ob ein Umstieg zu Open Access bei Zeitschriften auf breiter Front sich positiv auf die Bibliotheken auswirken würde, ist fraglich. Wenn sich die bisherige Publikationspraxis nicht verändert und die Verlage nur ihre bisher durch Zeitschriftenabonnements erzielten Einnahmen einfach auf Autorengebühren umwälzen, würden diese Gebühren mit Sicherheit aus den Bibliotheksetats entnommen. Zudem verlören die Bibliotheken eine weitere Aufgabe, da es dann ja keine Zeitschriftenabonnements mehr zu verwalten gäbe. Für die Bibliotheken attraktiver scheint das Open Access Modell im Bereich der Monographien, insbesondere der Hochschulschriften. Hier besteht durch den freien Internetzugriff ein Einsparpotential im Bereich Magazinflächen und Handlingkosten für die Printexemplare. Open Access Modelle im Bereich der Monographien werden bei den neueren Universitätsverlagen erfolgreich umgesetzt, zum Beispiel in Hamburg, Karlsruhe und Göttingen, die alle auf ein hybrides Publikationsmodell Open Access plus Print-on-demand setzen.

#### Digitalisierung - Papier ins Netz

Lag der Schwerpunkt bei der Digitalisierung von Bibliotheksbeständen in den 1990er-Jahren noch auf ausgewählten Schauobjekten wie der Gutenbergbibel (<http://www.gutenbergdigital.de>) oder den Handschriften der Bibliotheca Palatina (<http://www.ub.uni-heidelberg.de/helios/digi/palatina-digital.html>), so ist die Digitalisierung von Büchern inzwischen ein Massengeschäft. Das technische Equipment ist erschwinglich geworden, Scanroboter sind heute in der Lage, Bücher und Zeitschriften in großem Maßstab und in hoher Qualität einzuscannen.

Die Aktivitäten sind vielseitig: Die großen Wissenschaftsverlage haben inzwischen den größeren Teil ihrer Zeitschriftenarchive digitalisiert, die kleineren Verlagshäuser beginnen nachzuziehen. Große Monographiensammlungen wie Early English Books, Early American Imprints oder Netlibrary hat die DFG über Nationallizenzen für die deutschen Hochschulen zugänglich gemacht. Auf regionaler oder nationaler Ebene sind in den letzten Jahren umfangreiche Archive entstanden wie Gallica (<http://gallica.bnf.fr/>), Projekt Runeberg

(<http://runeberg.org/>) oder die Biblioteca digital Dioscórides der Biblioteca della universidad complutense de Madrid (<http://www.ucm.es/BUCM/buscar/6078.php>). Erste Versuche werden gemacht, die überall entstehenden, kleinen und großen Archive nachzuweisen, wie beim Zentralen Verzeichnis Digitalisierter Drucke (<http://zvdd.de>) oder beim BAM-Portal (<http://www.bam-portal.de/>).

Durch die zunehmende Digitalisierung älterer Bestände wächst auch der Druck auf die Verleger, auch die nicht mehr im Handel erhältlichen, noch urheberrechtsgeschützten Werke des 20. Jahrhunderts über Internet zugänglich zu machen bzw. deren Digitalisierung zu erlauben. Geschieht dies nicht, so bleibt die Literatur der letzten 100 Jahre im "Schwarzen Loch" des Urheberrechts gefangen und ist für die elektronische Informationsversorgung verloren. Google ist mit seinem Bibliotheksprogramm dabei, in Kooperation mit derzeit 19 Bibliotheken weltweit die gesamte Literaturproduktion des 19. Jahrhunderts frei im Internet über GoogleBooks zugänglich zu machen (<http://books.google.com/googlebooks/partners.html>). Google zeigt dabei, dass dies mit vertretbarem Aufwand und zu relativ geringen Kosten möglich ist. In der Bayrischen Staatsbibliothek schätzt man den Wert der geleisteten Digitalisierung - immerhin rund eine Million Bände - auf ca. 40 Millionen Euro. Sieht man diese Zahl im Vergleich zu den Kosten repräsentativer Bibliotheksneu- und Umbauten - den Umbau des Lesesaals der Staatsbibliothek unter den Linden beziffert man auf über 400 Millionen Euro - so rückt die Konversion der gesamten wissenschaftlichen und kommerziellen Buch- und Zeitschriftenproduktion in greifbare Nähe. Auch die DFG treibt die Digitalisierung in Bibliotheken weiter voran und wird für diesen Zweck in den nächsten Jahren einiges an Mitteln bereitstellen.

Für die Bibliotheken ist die Entwicklung eher zwiespältig, die massenhafte Digitalisierung erleichtert zwar den Nutzern erheblich den Zugang zur Literatur, doch wird sie letztlich dazu führen, dass die Bibliotheken ihrer bislang unbestrittenen Oberhoheit über den Zugang zur älteren Literatur verlustig werden. Es sei denn, den Verlagen gelingt es, die urheberrechtlichen Barrieren so zu erhöhen, dass Bibliotheksprogramme wie das von Google so eingeschränkt werden, dass es unbenutzbar wird.

Erschließung im Internetzeitalter - Fehlanzeige?

Interessanterweise hinkt die Erschließung der eigentlichen Digitalisierung deutlich hinterher. Eine inhaltliche Erschließung bieten nur ganz wenige Repositorien an, meist ist jede digitale Sammlung in einem anderen proprietären System eingebettet. Von einem einheitlichen Zugriff auf die Digitalisate und einem sachlichen Zugang ist man noch weit entfernt. Maßnahmen wie die Interaktion der Repositorien und der Austausch der Erschließungsdaten wären erste Schritte, um bei der rasant wachsenden Menge an Digitalisaten den Überblick zu behalten.

An den Datenformaten kann es nicht liegen; mit dem Dublin Core Datenset und der Verpackung dieser Daten in XML sind die technischen Voraussetzungen für den Transport von Erschließungsdaten erfüllt. Ein Hindernis ist sicherlich das schon im Zeitalter der Printmedien und Zettelkataloge beklagte Fehlen eines weltweit gültigen universellen Erschließungssystems. Bedauerlicherweise ist keine von der Library Community getragene Lösung des Problems in Sicht.

Wenig nachhaltig waren bislang auch alle Bemühungen, Internetquellen systematisch zu erschließen. Dies ist umso erstaunlicher, als doch gerade das Internet grundsätzlich die

Möglichkeit bietet, durch verteiltes kooperatives Arbeiten ohne allzu große Belastungen für einzelne voranzukommen. Obgleich immer wieder entsprechende Projekte gestartet wurden, ist daraus kein nachhaltiger Dienst entstanden. Wie die von Thomas Hilberer an der UB Düsseldorf initiierte und einst viel genutzte Düsseldorfer Virtuelle Bibliothek sind die allermeisten systematischen Sammlungen von Links und Internetdokumenten heute von geringer Bedeutung. Dies gilt allerdings nicht nur für die bibliothekarischen Projekte, sondern auch für andere umfassende systematische Linksammlungen. Die ehemals recht bedeutsamen Internetkataloge wie Dmoz (<http://www.dmoz.org/>), The WWW Virtual Library (<http://vlib.org/>), Deutsche Datenquellen (<http://www.uni-karlsruhe.de/Outerspace/VirtualLibrary/index.de.html>) führen heute im Vergleich zu Google ein Schattendasein. Dem Anschein nach befriedigt die bei Suchmaschinen benutzte Technologie die Informationsbedürfnisse der Nutzer deutlich besser als die manuell erstellten und gepflegten Internetkataloge. Wahrscheinlich wird auch das Erschließungsproblem durch einen rein technischen Ansatz gelöst werden. Entweder durch Integration der zahlreichen vorhandenen (Fach)Systematiken in semantische Netze oder durch die Analyse der Nutzung der Dokumente und daraus abgeleiteten fachlichen Clustern. Leider zeigen nur wenige Bibliotheken Ansatzpunkte, um zu diesen Entwicklungen aktive Beiträge zu leisten.

Die große Herausforderung für die Bibliothek wird in Zukunft sein, eine neue Aufgabe bei der Archivierung von elektronischen Dokumenten und Daten zu finden.

Suchmaschinen - ordnen die Information im Netz und machen sie zugänglich und nutzbar

Neben der Digitalisierung und der Verbreitung der Digitalisate über das Internet war die Erschließung der Internetquellen durch Suchmaschinen sicherlich die Entwicklung, die die Bibliothek am nachhaltigsten betrifft. 1990 wurde an der McGill Universität in Montreal der Dienst Archie entwickelt, um Informationen über Dateien auf FTP-Servern automatisch zu sammeln und den Benutzern aufbereitet zur Verfügung zu stellen. Archie war damit ein Vorgänger der Internetsuchmaschinen und für mehrere Jahre einer der am häufigsten genutzten Internet-Dienste. Den nächsten Fortschritt bei der Internetsuche brachte Veronica (Very Easy Rodent-Oriented Netwide Index to Computerized Archives), ein Dienst, der alle an der University of Minnesota gemeldeten Gopherserver indexierte. Gopher war dort 1991 als graphische Oberfläche für über das Web zugängliche Verzeichnisbäume entstanden. Veronica ermöglichte eine Suche über den monatlich aktualisierten Index und war damit der direkte Vorläufer der Suchmaschinen.

Die erste Suchmaschine wie wir sie heute kennen war der World Wide Web Wanderer, der von dem MIT-Studenten Mathew Gray 1993 für die Indexierung des gerade neu entstehenden World Wide Web programmiert wurde. Der Wanderer durchsuchte und indexierte zweimal pro Jahr das Netz. Das System wurde von Michael L. Mauldin weiterentwickelt, der später seinen eigenen Dienst Lycos anbot.

Im Dezember kamen die Suchmaschinen Jumpstation und Word Wide Web Worm hinzu, 1994 folgte WebCrawler und Lycos. Diese Suchmaschinen benutzen bei der Trefferanzeige erstmals Rankingalgorithmen für die Sortierung der Treffer nach Wichtigkeit.

# WWW - WORLD WIDE WEB WORM

Please note our new address: <http://www.cs.colorado.edu/www>



[Best of the Web '94](#) - Best Navigational Aid. [Oliver McBryan](#)

Serving 3,000,000 URL's to 2,000,000 folks/month.

[Instructions](#), [Definitions](#), [Examples](#), [Failures](#), [Register](#), [WWW Paper](#).

[VOTELINK](#) -- Check out this week's \*hot\* issues and VOTE.

1. Search all URL references

a. AND - match all keywords  
b. OR - match any keyword

Keywords:

Start Search

5 matches

1 match  
5 matches  
50 matches  
500 matches  
5000 matches

WordWideWeb Worm

Im selben Jahr entstand auch Yahoo, ursprünglich kein Suchmaschinenroboter, sondern eine manuell erstellte Linkliste der beiden Gründer David Filo und Jerry Yang in Stanford. Als die Liste immer länger und unhandlicher wurde, teilte man sie in verschiedene Kategorien und Unterkategorien auf. Diesem Grundkonzept folgt Yahoo bis heute. 1995 erschienen mit Infoseek und Altavista die ersten Suchmaschinen, die nicht in Universitäten, sondern von kommerziellen Firmen entwickelt wurden. Es folgten dann zahlreiche andere, inzwischen in Vergessenheit geratene Suchmaschinen wie Excite, Infoseek, Inktomi, Open Text und Northern Light, HotBot, die sich untereinander mehr oder weniger unterschieden.

Die größten Suchmaschinen (ohne Yahoo) waren den Statistiken von Greg Notess zufolge von 1996 bis 1999 Altavista, Northern Light und Hotbot

(<http://www.searchengineshowdown.com/statistics/size.shtml>). 1999 ging dann die Suchmaschine Fast (AllTheWeb) in Führung, doch schon 2001 übernahm Google den Spitzenplatz. Daran hat sich bis heute wenig geändert, Google liegt mit großem Abstand vorne, gefolgt von Yahoo und MSN Search. In Deutschland fallen 90 Prozent der Suchmaschinennutzung auf Google ([www.webhits.de/deutsch/index.shtml?webstats.html](http://www.webhits.de/deutsch/index.shtml?webstats.html)).

Ergänzt wurde der Suchmaschinenzoo durch eine bis zum Jahr 2000 wachsende Anzahl von Meta-Suchmaschinen wie MetaGer, MetaCrawler, Dog Pile, Metafind, die vor allem die geringe Erschließungsbreite und die schlechten Antwortzeiten der primären Suchdienste ausgleichen konnten. Durch die Dominanz von Google sind Metasuchmaschinen heute fast bedeutungslos geworden. Google ermöglichte vor allem eine effiziente und zuverlässige Suche, lieferte eine übersichtliche Trefferliste und war von Anfang an unerreicht schnell. Der große Erfolg von Google lässt sich auch darauf zurückführen, dass die Suchmaschine auf die Internetsuche konzentrierte und nicht dem zu Ende der 1990er-Jahre herrschenden Trend zum Portal folgte. Ironischerweise ist Google inzwischen dabei, genau diese Portalfunktion

auszufüllen. Mit einem vielfältigen Portfolio von Zusatzdiensten, die um die Internetsuche entstanden, kann Google heute fast alle Bedürfnisse des Internetnutzers befriedigen, von der Textverarbeitung über E-Mail und Bildbearbeitung bis hin zur Urlaubsplanung.

Suchmaschinen wurden von den Bibliotheken bis vor kurzem meist entweder ignoriert und belächelt, und Bibliothekare waren bei der Entwicklung von Suchmaschinen so gut wie gar nicht beteiligt, ein merkwürdiger Umstand, wenn man bedenkt, dass Sammeln, Erschließen und Bereitstellen von Informationen die Kerngebiete dieses Berufsstandes ausmachen. Erst mit der Popularisierung der Internetsuche durch Google in den letzten Jahren und dem Auftritt von Google Books im Jahr 2004 reifte die Erkenntnis, dass in den Suchmaschinen eine starke Konkurrenz zu den Bibliotheken erwachsen ist.

Sie bedrohen deren tradierte Rolle als Zugangspunkt zum Weltwissen, wie es das Motto von Google "to organize the world's information and make it universally accessible and useful" (<http://www.google.com/corporate/>) unübersehbar zum Ausdruck bringt. Google und seine Konkurrenten Yahoo und Microsoft werden mit diesem Anspruch, das Wissen der Welt zu präsentieren, nicht vor den Pforten der altherwürdigen Bibliotheken haltmachen. Die Digitalisierung von Millionen von Büchern, die Google in derzeit 19 Bibliotheken vornimmt, wird die Bibliothekswelt nachhaltig verändern. Mindestens so sehr wie die Einführung des Internets.

---

**Autor**

**Dr. Michael W. Mönnich**

**Universitätsbibliothek Karlsruhe**

Kaiserstr. 12

76049 Karlsruhe

[Michael.Moennich@ubka.uni-karlsruhe.de](mailto:Michael.Moennich@ubka.uni-karlsruhe.de)

